REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*
31-08-2009

2. REPORT TYPE
FINAL

3. DATES COVERED *(From - To)*
15/09/2006 - 31/08/2009

4. TITLE AND SUBTITLE
Architecting Science:

Practical Tools for Architecting Flexible Systems

5a. CONTRACT NUMBER

5b. GRANT NUMBER
FA9550-0601-0550

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Daniel E. Hastings and Mary L. Cummings

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Massachusetts Institute of Technology
77 Massachusetts Avenue
Cambridge, MA 02139

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AFOSR
875 N Randolph St.
Arlington, VA 22203 David.Luginbuhl@afosr.af.mil
Dr David Luginbuhl/RSL

10. SPONSOR/MONITOR'S ACRONYM(S)
AFOSR

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Distribution A: Approved for Public Release

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Three areas of architecting science were investigated: (1) Tradespace exploration is a conceptual design tool used to compare thousands of designs. Flexibility, the ability to dynamically change systems to mitigate risk or leverage opportunity, is difficult to assess within static tradespaces. A metric to identify valuably flexible designs in a tradespace, Value Weighted Filtered Outdegree, was developed and applied to a satellite radar and ORS case, and shown to identify designs that would not have been identified with existing tradespace metrics. (2) One cited mechanism for accommodating uncertainty in design is to embed flexibility. A new framework was developed in the research to aid the design of a flexible system. Change propagation analysis was extended for analysis of systems with heterogeneous relationships between components, using filtered outdegree analysis to quantify flexibility. (3) Decision makers utilize multiple criteria and many levels of reasoning to understand complex multivariate data. They can be aided by direct-perception decision-support systems of configural rather than separable displays. An experiment compared performance across different levels of reasoning, finding that configural display promoted better performance and more efficient eye fixation patterns at the highest level of reasoning than the separable display, and was the subjective tool of choice.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF: UNCLASSIFIED			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Daniel E. Hastings
					19b. TELEPHONE NUMBER *(include area code)* 617-253-6056

20100426218

AFOSR Final Project Report "Architecting Science: Practical Tools for Architecting Flexible Systems"

Professor Daniel E. Hastings, Principal Investigator
Professor Mary L. Cummings, Co-Principal Investigator

Massachusetts Institute of Technology

August 31, 2009

The AFOSR grant "Architecting Science: Practical Tools for Architecting Flexible Systems" (Grant No. FA9550-0601-0550). The project was funded for a period of 9/15/06 to 8/31/09. This report summarize the significant research outcomes of the project in three sub-topic areas.

Executive Summary

Systems architectures are growing increasingly complex as technology evolves and systems of systems are developed, comprised of legacy systems with new interfaces. Senior leaders in the Air Force have noted that the Air Force is creating more of these complex architectures. Based on past experience, these systems will remain in service for a long time, often significantly past their nominal design lives. Early decision making is critically important, but difficult due to level of uncertainty inherent in this phase of programs. These factors drive the need for advanced constructs, metrics, and decision-aiding tools in the architectural phase of system programs, and motivated this research.

The typical practice in making architectural choices by senior decision makers is as follows: A high level need or set of needs is identified (for example, global positioning and timing to mobile forces). Various model architectures which meet the initially expressed needs are developed (sometimes with some measure of user utility) and performance issues are traded against each other. In the last decade, cost has also been traded as an independent variable (the so-called Cost As an Independent Variable (CAIV) trades). The range of options considered for the architectural trades are often shown in a tradespace analysis where utility is plotted against cost. However, often the only part of the tradespace explored in detail is the local space around selected point designs. Senior decision makers, guided by experience and heuristics, then make choices based on the information presented that a certain set of architectures is optimal for meeting the initially expressed set of needs. Out of the wide variety of uncertainties that a new system faces, only a limited set of technical and development risks (e.g. those that can be handled by margins, redundancies and risk management plans) are considered. The "unknown-unknowns" that might await the system in the future, especially those concerning changing needs and threats, and the "upside" to uncertainty, where the system may prove much MORE useful than originally envisioned, are considered qualitatively if at all. Recognition of this issue has led to the call for "robust," "flexible" and/or "evolvable" architectures.

The state of the practice has a number of well-recognized flaws. The first is that the stakeholders may not express all their needs. This is not because of malfeasance but because stakeholders may not know what they all are or may not realize the real capabilities of a system. The second is that the process produces architectures which are very brittle to the inevitable changes in funding, technological changes, new stakeholders etc.[1] In other words, the current practice does a very poor job in giving the decision makers choices that will be robust to future uncertainty. Thirdly, there is only a limited set of heuristics to guide decision makers on where to embed flexibility in the system architecture. For example, one of the heuristics at AF SMC is to over design the processors on their satellites by 10-20% relative to the initial processing load. They use this based on past experience of the growth of processing need on their satellites. Of course the problem is that the over design of the processor cascades into the thermal management system, the power system etc. Therefore the flexibility has a real cost associated with it for a future uncertain benefit. Thus the state of practice does not understand how and when the options for flexibility should be embedded in a real system. Finally, it is very hard to convey appropriate tradeoffs and choices when there are multiple decision makers who all have important stakes in a system. Coupling of the architectural choices to the human decision making process needs to be substantially improved.

In the research, three critical areas were addressed that aimed at improving the state of the art as well as the state of practice. MIT concentrated on the architecting of robust/flexible/evolvable systems in order to handle and exploit the uncertainty inevitable when designing complex systems which will have long lifetimes, be asked to perform in a variety of systems-of-systems, and face uncertain missions, threats, and environments. The three areas of study included 1) how to quantify "flexibility" and calculate its potential value on complex architecture tradespaces; 2) how to identify metrics of flexibility in architectures, and how to include options to deal with uncertainty, within a given system architecture and 3) how to enhance decision support systems through display of complex information in a more intuitive, principled format.

Three detailed white papers included in this report describe each of the following significant outcomes of the research:

1. Development of a new tradespace metric, Value Weighted Filtered Outdegree (VWFO), for identifying valuably flexible designs with application on a satellite radar and Operationally Responsive Space (ORS) case.
2. Implementation of a Coupled-DSM for a Micro Air Vehicle (MAV) with change scenario definitions, with preliminary results for formulating real options based on change propagation analysis
3. A new configural display, Fan Visualization (FanVis) [2] that takes system acquisition tradespace data, and using emergent features, naturally maps data for the decision maker.

[1] Recent theoretical results suggest that the more optimized a system is to *anticipated* environments and threats, the more vulnerable it will be to unanticipated changes. This is precisely the situation that many defense aerospace systems find themselves in: highly optimized for missions that may be changing or outdated. See Carlson, J. M. and Doyle, J, "Highly Optimized Tolerance: Robustness and Design in Complex Systems," Physical Review Letters, Vol. 84, No. 11, March 2000, pp 2529-2532.

[2] A technology disclosure was made to the MIT Technology Licensing Office for the software developed in this project. Negotiations are currently underway for a private company to license the software developed under the portion of this grant.

The contributions of the research have furthered the evolution of system architecting from an art to a science. Further background information and related research results can be found on the MIT websites for the SEAri (http://seari.mit.edu) and Human and Automations Lab (http://halab.mit.edu).

The research has resulted in several theses, and journal and conference papers.

Publications

Theses

Massie, A. E., (2009), Designing a Graphical Decision Support Tool to Improve System Acquisition Decision, S. M. Thesis, MIT Aeronautics and Astronautics, Cambridge, MA (http://halab.mit.edu)

Viscito, L., Quantifying Flexibility in the Operationally Responsive Space Paradigm, Master of Science Thesis, Aeronautics and Astronautics, MIT, June 2009 (http://seari.mit.edu)

Wilds, J.M., A Methodology for Identifying Flexible Design Opportunities, Master of Science Thesis, Technology and Policy and Aeronautics and Astronautics, MIT, September 2008 (http://seari.mit.edu)

Refereed Journals

Massie, A. and Cummings, M., Designing a Perception-Based Decision Support System to Improve System Acquisition Decisions, Decision Support Systems, in review

Conference Papers

Bruni, S., Marquez, J.J., Brzezinski, A., Nehme, C.,. and Boussemart, Y., Introducing a Human-Automation Collaboration Taxonomy (HACT) in Command and Control Decision-Support Systems, 12th International Command and Control Research and Technology Symposium, Newport, RI, June, 2007 (http://halab.mit.edu)

Massie, A., Kopylov, I., and M.L. Cummings, Supporting System Acquisition Decisions through Ecological Perception" 47th AIAA Aerospace Sciences Meeting, Orlando, FL 2009 (http://halab.mit.edu

Ross, A.M., Rhodes, D.H., and Hastings, D.E., "Defining Changeability: Reconciling Flexibility, Adaptability, Scalability and Robustness for Maintaining Lifecycle Value," INCOSE International Symposium 2007, San Diego, CA, June 2007 (http://seari.mit.edu)

Ross, A.M. and Hastings, D.E., "Assessing Changeability in Aerospace Systems Architecting and Design Using Dynamic Multi-Attribute Tradespace Exploration," AIAA Space 2006, San Jose, CA, September 2006 (http://seari.mit.edu)

Shah, N.B., Viscito, L., Wilds, J.M., Ross, A.M., and Hastings, D.E., "Quantifying Flexibility for Architecting Changeable Systems," 6th Conference on Systems Engineering Research, Los Angeles, CA, April 2008 (http://seari.mit.edu)

Viscito, L., and Ross, A.M., "Quantifying Flexibility in Tradespace Exploration: Value-Weighted Filtered Outdegree," AIAA Space 2009, Pasadena, CA, September 2009 (http://seari.mit.edu)
Viscito, L., Chattopadhyay, D., and Ross, A.M., "Combining Pareto Trace with Filtered Outdegree for Identifying Valuably Flexible Systems," 7th Conference on Systems Engineering Research, Loughborough University, UK, April 2009(http://seari.mit.edu)
Viscito, L., Richards, M.G., and Ross, A.M., "Assessing the Value Proposition for Operationally Responsive Space," AIAA Space 2008, San Diego, CA, September 2008(http://seari.mit.edu)
Wilds, J., Bartolomei, J.E., de Neufville, R., and Hastings, D.E., "Real Options "In" a Micro Air Vehicle," 5th Conference on Systems Engineering Research, Hoboken, NJ, March 2007(http://seari.mit.edu)

CONTRIBUTORS

Professor Daniel E. Hastings	Principal Investigator	Task 1 and Task 2
Professor Mary L. Cummings	Co-Principal Investigator	Task 3
Dr. Adam M. Ross	Research Scientist	Task 1 and 2
Dr. Donna H. Rhodes	Principal Research Scientist	Task 1 and 2
Anna Massie Bethke	Graduate Research Assistant	Task 3
Amy Brzezinski	Graduate Research Assistant	Task 3
Nirav Shah	Graduate Research Assistant	Task 2
Lauren Viscito	Graduate Research Assistant	Task 1
Jennifer Wilds	Graduate Research Assistant	Task 2
Kevin Boehm	Undergraduate Research Student	Task 3
Sean Jameson	Undergraduate Research Student	Task 3
Igor Kopylov	Undergraduate Research Student	Task 3

Quantifying Flexibility in the Operationally Responsive Space Paradigm

Lauren Viscito
Advisors: A. M. Ross and D. H. Rhodes,

Abstract

Designing complex space systems that will deliver value in the presence of an uncertain future is difficult. As space system lifetimes are now measured in decades, the systems face increased risk from uncertain future contexts. Tradespace exploration increases the designer's system knowledge during conceptual design and with dynamic analysis can predict the system's behavior in many possible future contexts. Designing flexible systems will allow mitigation of risk from uncertain future contexts and the opportunity to deliver more value than anticipated by the designers.

Flexibility is a dynamic property of a system that allows it to take advantage of emergent opportunity and to mitigate risk by enabling the system to respond to changing contexts in order to retain or increase usefulness to system stakeholders over time. Value Weighted Filtered Outdegree is introduced as a metric for identifying valuably flexible systems in tradespace studies in order to improve decision making during the conceptual design phase. Dynamic Multi-Attribute Tradespace Exploration (Dynamic MATE) is used as the basic tradespace exploration method for Value Weighted Filtered Outdegree, and applies decision theory to computer simulation of thousands of system designs, across hundreds of unique future contexts. Epoch-Era Analysis is used to parameterize future contexts for dynamic analysis of the designs' performance. Although dominated in static analysis, flexible designs are valuable in the presence of changing contexts.

1. Introduction

Since the launch of the first artificial satellite, Sputnik, on October 4, 1957 the uses and applications of satellites have become entwined in national security, commercial ventures, and everyday households. Timing signals from Global Positioning Systems constellations enable ATM transactions, voice and telecommunications are transmitted almost instantly around the world, and military operations in far-flung Areas of Interest (AOIs) are supported in real time from locations within the nation. Satellites have become an accepted and essential component of doing business in the modern world, and as engineers look to the future the demand for space-based capability is likely to increase.

2. Problem Formulation

As the expected lifetime of a satellite system increased from years to decades, the need to keep the system relevant became more pronounced. Flexibility is a dynamic property of a system that allows the system to take advantage of emergent opportunity and to mitigate risk by enabling the system to respond to changing contexts in order to retain or increase usefulness to system stakeholders over time (Ross et al., 2008; Viscito and Ross, 2009). A space system that is flexible will deliver more value over these changes than a non-flexible system. Increasing system costs and operational lifetimes have driven research in recent years to develop understanding of how flexibility can be incorporated into systems in the conceptual design phase (Saleh et al., 2008). However, the concept of flexibility is still not mature, and designing for flexibility would be aided by a way to compare the flexibility of systems during conceptual design.

2.A. Problem Statement

Space systems are essential to the national security of the United States of America and many other nations. Typically, designers would consider objectives such as cost, mass, volume, and capability, and the satellite that minimized mass, volume and cost at the greatest capability would be the optimal design. However, the lifetime of the system can often be measured in decades. Over that period of time, the likelihood of the environment and user requirements changing is very high. If the context does change, it is possible the system will no longer deliver acceptable levels of value. Value is a subjective measure of benefit from a bundle of consequences that is specified by a stakeholder (Keeney and Raiffa, 1993).

2.B. Research Questions

This research seeks to answer two questions.

- What is an objective, repeatable metric that incorporates design utility and flexibility?
- Does a modular architecture have more flexibility than a legacy architecture for an electro-optical imaging Operationally Responsive Space mission with changing user preferences?

In answering the first question, the process of developing the metric will be described, and then the metric is demonstrated in a case application. Dynamic MATE (Ross, 2003, 2006), and Epoch-Era Analysis (Roberts et al., 2009), are explained within the implementation framework Responsive Systems Comparison (RSC) (Ross et al., 2008). The extensions for calculating the metric are described using the information generated during RSC.

The second question deals with applying the new metric to a current topic and exercising the new metric. A case application about Operationally Responsive Space (ORS) (Cerbowski and Raymond, 2005; Department of Defense, 2007; Fram, 2007) will compare the flexibility of two architectures for ORS: legacy or 'custom', and modular.

3. *Value Weighted Filtered Outdegree*

Because the value of flexibility is only realized in the presence of uncertainty, the designer needs to have a possible future era in which to asses the design in the tradespace. Value Weighted Filtered Outdegree is defined as:

$$VWFO_i^k = \frac{1}{N-1}\sum_{j=1}^{N-1}\left[sign\left(u_j^{k+1} - u_i^{k+1}\right) * Arc_{i,j}^k\right]$$ where

N is the number of designs considered
k is the current epoch
$k+1$ is the next epoch in the era
i is the design under consideration
j is the destination design
u_i^{k+1} is the utility of design i in the $k+1$ epoch
u_j^{k+1} is the utility of design j in the $k+1$ epoch
$Arc_{i,j}^k$ is the transition matrix with local value indicating an arc from design i to design j in epoch k

The analyst can choose to look at the VWFO of an entire design space, in which case N is the same as the total number of designs in a the tradespace study. Alternatively, a smaller subset of designs can be chosen, and examined in great detail. VWFO uses the direction change in utility to determine if a particular transition is 'good', which occurs when the design transitions to a design of higher utility. By summing both the positive and negative transitions, the designer can see designs that are valuably flexible, (the design with positive VWFO), and the designs that are changeable but are carrying 'dead weight' (the design with negative VWFO).

3.A. Satellite Radar System Case Application

The Satellite Radar System is an extensive modeling effort performed by several students. The model follows the RSC processes, which are described in Ross et al. (2008).
This metric captures the utility difference in the destination designs and is dependent on how many transitions are available. The intent of the metric is to act as a screening heuristic, and it is left to the decision maker to make a final call on the value of the design.

Designs with high magnitudes of VWFO may be more valuably flexible than others. Designs that have positive VWFO are able to transition to destination designs that have higher net utility. Unlike choosing designs based solely on high NPT or high FOD, VWFO can identify designs that are valuable and flexible. VWFO takes into account the value of the change (the utility change direction), the changeability of the design (transition arcs), and the context changes (era progression).

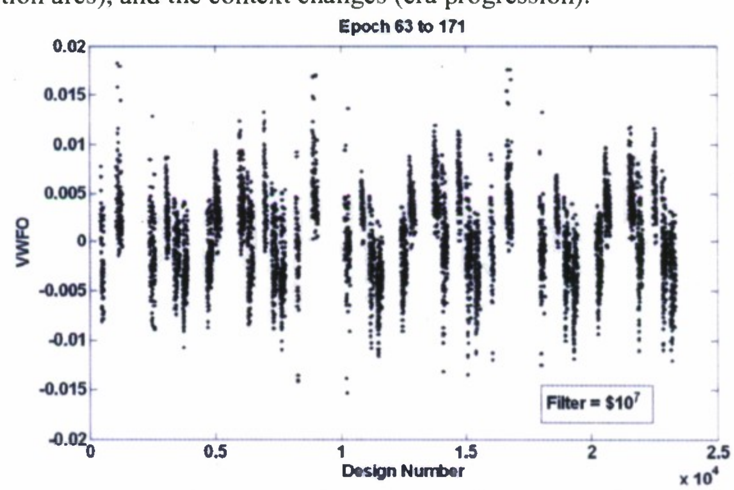

Figure 1: Value Weighted Filtered Outdegree for Epoch 63-171

Figure 1 plots the VWFO for an entire tradespace by design number. The striations in the space are caused by the discrete enumeration of the design space.
The SRS case application revealed several interesting aspects of VWFO. Several designs had high VWFO, and one way for a designer to determine which design to analyze further is to use Figure 2, which indicates designs with high VWFO and high origin design utility. By having positive VWFO and high starting utility, these designs have more transitions to other high utility designs.

There are several limitations to VWFO. The first is that the results obtained from any study of this nature will depend on the transition rules chosen. If the designers do not specify transition rules that are useful during context changes, no designs will be identified. In addition, it is also dependent on the order of epochs in the era, which determines which designs are valid for transition. Essentially, if a design in invalid in either epoch, it appears as invalid in both. Another problem with the metric occurs when the VWFO of a design is zero.

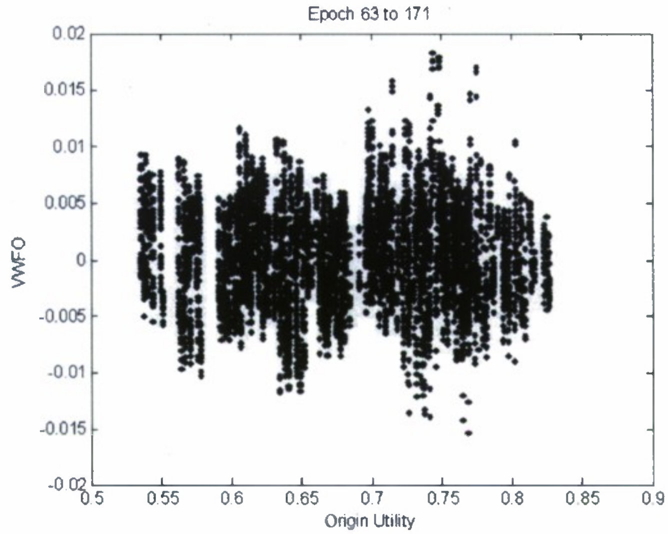

Figure 2: Total Utility of Origin Design with VWFO to Check for Linear Correlation

The designer does not know, without further analysis, if design has zero VWFO because it is an invalid design in one of the epochs, or because the net utility change is zero. The metric is also dependant on the tradespace sampling strategy used by the designer. If the design space has many designs in one area of the design space, which causes the FOD of those designs to increase, it is likely that the VWFO of the design will increase disproportionately as well.

3.B. Operationally Responsive Space Case Application

The ORS case application modeled a small electro-optical imaging spacecraft (Richards et al., 2008; Viscito et al., 2009), constructing a legacy or modular spacecraft system for the designs. The modular designs had an expansion option available in the event that changes were required during the early phases of the spacecraft. Figure 3 shows VWFO plotted by design number for each of the epoch transitions.

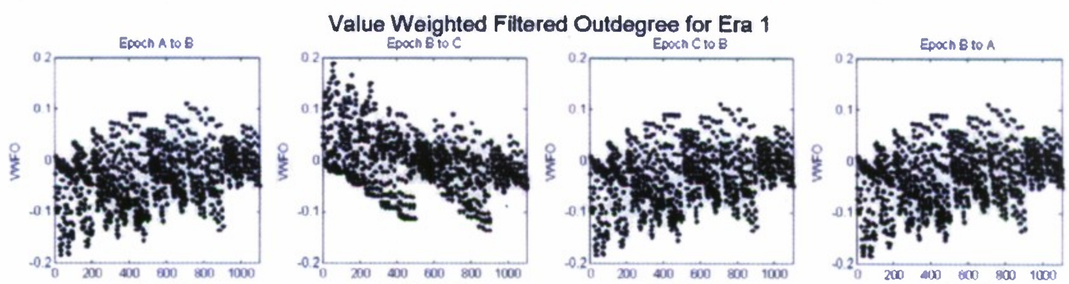

Figure 3: Value Weighted Filtered Outdegree for ORS Era 1

The designs with the highest VWFO are the modular architecture designs. Interestingly, these designs are not the highest utility designs, as indicated by the Pareto sets , nor the highest Filtered Outdegree designs, reinforcing the intuition gained from the previous case study that there is a sweet spot for flexibility somewhere off the Pareto Front. Modular designs are not required to be on the Pareto Front, as the value of the architecture manifests when stakeholders' needs change. The ease with which a modular spacecraft can be built to meet these new needs is more than that of the legacy systems. While the legacy systems may be more optimal in the traditional sense, under the paradigm shift where responsiveness is more important than perfection, a modular architecture may be beneficial.

The differences between legacy and modular architectures, as modeled, are shown in the static tradespace and in the VWFO analysis. While traditional metrics such as Pareto Trace and Filtered Outdegree identified two sets of designs, VWFO indicated a design set separate from those identified earlier. As the choice for ORS systems, a modular architecture looks like a good choice. Designs with modular architectures have higher VWFO than many of the legacy architectures.

A note of caution, however, as the result will be very sensitive to the actual ability of the modular architecture to achieve the schedules and costs estimated in this model. If the modular architecture is used to advantage, meaning that many spacecraft are built taking advantage of the economies of scale available, then the results may hold true. However, the experimental nature of the ORS paradigm should signal that many of the assumptions made in this model may prove to be invalid.

4. Conclusions

While previous metrics measured the passive value robustness of a design (Normalized Pareto Trace) or the changeability of a design (Filtered Outdegree), neither has addressed the how to identify valuable changes. Flexibility, much like optimization, is only optimal with respect to a defined objective, and therefore saying a design is flexible begs the question of what contexts it is flexible to. Changeable designs have many ways to transition to other designs, but these designs are only valuable if the transition can mitigate utility loss, either in the near or long term. Interest in flexibility as a risk mitigation strategy has increased, so it has become necessary to create a metric for identifying valuably flexible designs. The operationalization of valuable flexibility from a general concept to a tradespace system property will aid decision makers and increase their ability to make the business case for including flexibility.

Flexibility is the dynamic property of a system that allows it to take advantage of emergent opportunity and to mitigate risk, by enabling the system to respond to changing contexts, in order to retain or increase usefulness to system stakeholders over time. As the lifetime of a system approaches decades long, the likelihood of encountering changing contexts increases, leading to the risk that the system will not be able to deliver enough value in the new context. This risk, along with the desire to take advantage of opportunities also presented by the changing contexts, has increased the desire to include flexible systems in conceptual design trade studies. However, identifying these flexible systems has often been a subjective and haphazard endeavor.

Dynamic MATE, a method using utility theory and computer simulation to aid decision making during front-end conceptual design, was used as the framework for developing a flexibility metric. Several existing MATE metrics, Normalized Pareto Trace and Filtered Outdegree, informed the new metric for flexibility: Value Weighted Filtered Outdegree. VWFO is objective and repeatable; it identifies the valuably flexible designs by combining changeability with the value as operationalized by utility changes. To incorporate the dimension of time, which is necessary for flexibility to be valuable, future scenarios were parameterized with Epoch-Era Analysis.

VWFO acts as a screening heuristic on a design space, identifying designs that are valuably flexible. This gives a decision maker several designs with which to focus the search for flexibility. When presented with the thousands of designs that may be generated in a front-end tradespace study, any decision maker may be quickly overwhelmed with the amount of data that is presented. As designers, distilling this much data looking for the 'best' designs can be a daunting task. By operationalizing the 'ility' of flexibility, the designers have been given a starting point.

VWFO is a good metric for capturing flexibility for three reasons.

- It is objective and repeatable.

- The assumptions and biases in the metric can be understood by examining the transition rules, the epochs and eras chosen, and the model assumptions.
- It identifies a subset of designs that are highly changeable and valuably flexible.

Static analysis of the ORS tradespaces revealed that in general, the modular architectures had more utility than the legacy architectures, for a higher cost in the short term, or for a constant utility, higher cost. While the legacy architectures had lower initial costs and tended to have higher Filtered Outdegree, many modular designs had higher VWFO. This would suggest that the modular architecture, for the given transition rules and modeling assumptions, does have more flexibility than the legacy architecture. Given the low confidence in the modular cost estimation models, a sensitivity analysis is necessary to see if this is actually the case, however this was outside the scope of the case application.

References

Cerbowski, A. K. and J. W. Raymond (2005). Operationally responsive space: A new defense business model. Parameters, US Army War College Quarterly Summer 2005, 10.

Department of Defense (2007, 17 April). The plan for operationally responsive space. Technical report, U.S. Government.

Fram, B. J. (2007). The case for operationally responsive space. In 5th Responsive Space Conference, Los Angeles, CA.

Keeney, R. L. and H. Raiffa (1993). Decisions with Multiple Objectives. Cambridge: Cambridge University Press.

Richards, M. G., L. Viscito, A. M. Ross, and D. E. Hastings (2008). Distinguishing attributes for the operationally responsive space paradigm. In Responsive Space 6 Conference, Los Angeles, CA.

Roberts, C. J., M. G. Richards, A. M. Ross, D. H. Rhodes, and D. E. Hastings (2009). Scenario panning in dynamic multi-attribute tradespace exploration. In IEEE International Systems Conference, Vancouver, Canada.

Ross, A. M. (2003). Multi-attribute tradespace exploration with concurrent design as a value-centric framework for space systems architecture and design. Master's thesis, Massachusetts Institute of Technology.

Ross, A. M. (2006). Managing Unarticulated Value: Changeability in Multi-Attribute Tradespace Exploration. Ph. D. thesis, Massachusetts Institute of Technology.

Ross, A. M., H. L. McManus, A. Long, M. G. Richards, D. H. Rhodes, and D. E. Hastings (2008). Responsive systems comparison method: Case study in assessing future designs in the presence of change. In AIAA Space 2008, San Diego, CA.

Ross, A. M., D. H. Rhodes, and D. E. Hastings (2008). Defining changeability: Reconciling flexibility, adaptability, scalability, modifiability, and robustness for maintaining lifecycle value. Systems Engineering 11 (3), 246–262.

Saleh, J. H., G. T. Mark, and N. C. Jordan (2008). Flexibility: a multi-disciplinary literature review and a research agenda for designing flexible engineering systems. Journal of Engineering Design.

Viscito, L., M. G. Richards, and A. M. Ross (2009). Assessing the value proposition for operationally responsive space. In AIAA Space 2009, San Diego, CA.

Viscito, L. and A. M. Ross (2009). Combining pareto trace and filtered outdegree as a metric for identifying valuably flexible designs. In 7th Conference on Systems Engineering Research, Loughborough, UK.

System and Component Level Flexibility of a Micro Air Vehicle: A Case Study in Flexible System Design using Change Propagation Analysis and Filtered Outdegree Methods

Jennifer M. Wilds and Nirav B. Shah

Advisor: Professor Daniel Hastings

Abstract

The design of complex systems often, if not always, occurs in a context that is uncertain --needs change, technology evolves, and resources are uncertain. Much recent work has focused on design of systems that are able to deliver high value over time despite uncertainty. One commonly cited mechanism for doing so is to embed flexibility into system design. The design of a flexible system is described in terms of a new framework. Change propagation analysis is extended to allow analysis of systems with heterogeneous relationships between system components. Filtered outdegree analysis is presented as a method for quantifying flexibility at the system level. This case example supplements the theory development in Shah et al (2008). Change propagation methods and filtered outdegree are used to consider the formulation of real options for flexible system design for a Micro Air Vehicle (MAV).

1. Background

The US military is developing a micro air vehicle to provided reconnaissance and surveillance at a greater standoff distance. A micro air vehicle (MAV) is a less-than-one-pound unmanned air system equipped with a visual sensor. Conceptual design of the system continues to progress and formal requirements are being developed. Customers are willing to accept a less than optimal initial design, but ultimately would like to acquire one that can be adapted to changing operational needs. In addition, enabling technologies are evolving at a rapid rate relative to the program's production cycle.

Uncertainties. Key uncertainties were identified to analyze the design of the MAV within this uncertain context. Operational uncertainties considered are the required range and endurance, while technological uncertainties include the availability of advanced sensors and of high energy density power supplies. Based on these uncertainties, several change scenarios were identified as likely future environments in which the system may be required to operate. Three such scenarios enumerated below are the focus of this case study.

1. A technological change in the payload to enable day/night operations (CS #1)
2. A change in the data transmission standoff requirement (CS #2)
3. A change in the endurance requirement to allow entry into a new market (CS#3)

Physical Description. MAVs contain three major components: the air vehicle, the ground station, and the operator control unit, which is a software application providing a graphical user interface. The complexity of the interactions between the three components is beyond the scope of this analysis, thus for simplicity the system analyzed in this paper will be restricted to only the air vehicle and the ground station components. The air vehicle will include all components within the physical airframe, including the airframe itself. The analysis will be further limited to consideration of hardware components to improve or maintain performance, rather than modifications to the software algorithms within the autopilot, data link or mission controller.

The airframe can be decomposed into a series of objects, which can be described in terms of geometric and mass properties. Figure 1 shows the various components of the MAV airframe. A physical model of the air vehicle design was developed using MS Excel© by the USAF Academy (Bartolomei 2005) and validated by

Air Force Research Laboratory, Munitions Directorate for a series of MAV platforms. The model accepts geometric and mass property inputs for components of the MAV to return performance objectives, such as endurance, range, and airspeed solutions. The model enables designers to quickly compute impacts to performance resulting from changes to the physical design.

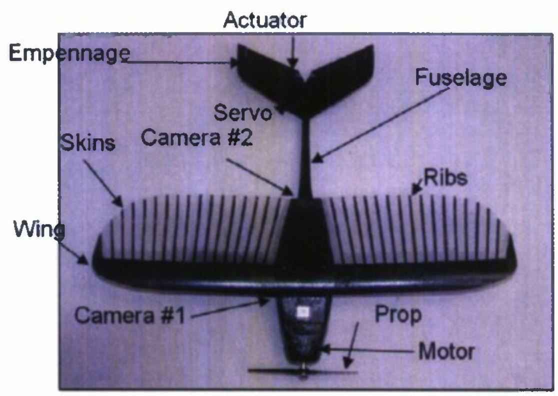

Figure 1. The Anatomy of a Micro Air Vehicle (Wilds et al 2007)

2. *Application of Change Propagation Analysis*

Clarkson et al. (2001) present a framework for analyzing the propagation of change throughout a system. Because change becomes more costly as the design matures due to integration efforts (i.e., a change to one part is more likely to affect multiple additional parts), it is advantageous to understand how the system will respond to future change and, if possible, create a design that is flexible to those changes. This paper will attempt to use the CPA method, along with the Change Propagation Index (CPI) introduced by Suh (2005) to identify candidates for embedded flexibility at the component level.

The Physical DSM. The first step of the analysis involves the creation of a DSM representing the MAV system. The DSM includes a physical decomposition of the ground station and the air vehicle to the component level, resulting in seventy-two nodes. Then, four types of relationships are recognized as existing between the physical components: power (electrical flows), data transmission (information flows), hardware interface (a spatial relationship indicating adjoining parts or physical connection), and "housing" (a geometric constraint relation indicating physical location). The result is a directed graph indicating the nodal relationships represented as edges. The matrix is sparsely populated; however the system as a whole is highly connected as a result of a tightly integrated system, as seen in Figure 2.

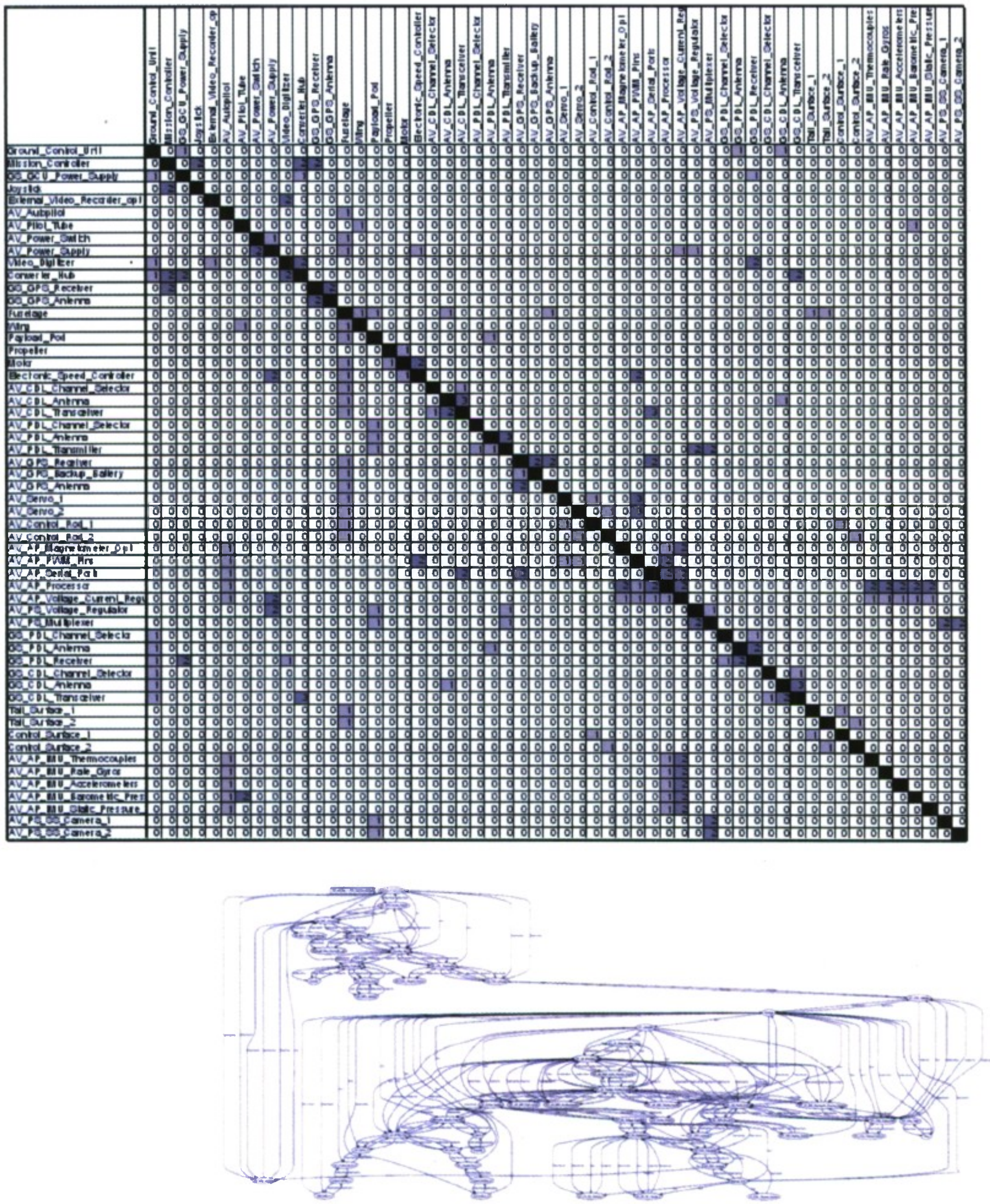

Figure 2. Design System Matrix (top) and Directed Graph (bottom) of the Micro Air Vehicle System

Change Scenarios. Change propagates when the tolerance margins of individual parameters are exceeded. (Eckert et al. 2004) Tolerance margins often include a design buffer, or contingency margin, which is used to absorb emergent changes that are not known at the time of design. Contingency margins are often decided based on primitive assessment of future uncertainties. Therefore, the first step in analyzing potential change propagation is identifying change scenarios based on the future uncertainties. This paper considers three change scenarios as previously defined. External uncertainties drive the internal changes within the system. Thus, change scenarios ask the question: "If component A is required to change due to resolved uncertainty, what other components will also change?" In this example, component A is the change initiator, or point where the change is introduced into the system. The question is answered based on the assessment of the perceived magnitude of change required as compared to the component's contingency margin. If the margin is exceeded, then change is propagated, however, if the margin is not exceeded, then change is absorbed. There may be multiple change initiators for a given change scenario. This occurs when the change is introduced into multiple components simultaneously.

In this case study, an interview with a Subject Matter Expert (SME) was conducted to determine the change initiators for each of the change scenarios. The SME was asked to consider which components within the system would likely be changed in direct response to the change scenario. For example, the first change scenario (CS #1) considers a change in technology that is most likely a newly available sensor suite. The SME indicated that the existing sensor suite is the primary change initiator in the system in response to a new or upgraded sensor. In a less intuitive scenario, for example the third change scenario, the objective/functional requirements flow down to the physical components is necessary to identify the appropriate change initiators. In the case of a change in the endurance requirement (CS #3), the SME considered the components which directly contributed to the endurance computation, leading to five possible change initiators. Table 1 displays the identified change initiators for each respective change scenario.

Table 1. Identified Change Initiators for Particular Change Scenarios

Change Scenario	Change Initiator(s)
Payload (CS#1)	▪ Payload Sensor Suite
Range (CS #2)	▪ Power Supply ▪ Comm Data Link Antennas ▪ Payload Data Link Antennas ▪ Payload Data Link ▪ Comm Data Link
Endurance (CS#3)	▪ Power Supply (AV) ▪ Propeller ▪ Motor ▪ Electronic Speed Controller (ESC) ▪ Wing

The SME was then asked to identify which of the relationship types (i.e., power, data transmission, etc.) are most directly affected in each change scenario. Because the DSM was limited initially to only four types of relationships, all four relationships were included in the analysis of each change scenario. However, in a more inclusive data set, filtering the connections that are not effected by the change scenario may reduce the computational complexity of the analysis.

Filtered Undirected Graphs. The next step of the analysis requires an algorithm that filters the DSM to include only the types of relationships affected in each change scenario. Recall that the DSM represents a directed connectivity graph. This graph may include relationship directions according to the system flows; however those directions may not be representative of the change flows depending on where the change is introduced in the system. Because changes can propagate both upstream and downstream from where the change initiator is located, the DSM must first be altered to reflect an undirected graph (i.e., the DSM is symmetric about the diagonal). Given a particular change scenario, the identified change initiators, and the respective relationship types affected by the change, the algorithm then searches the DSM for all defined connections between nodes matching the relationship type. The result is an undirected sub-graph including only the nodes within the new filtered network for each relationship type. For example, CS #1 has one change initiator (Payload Sensor Suite) and four relationship types (power, data transmission, hardware interface, houses). Therefore, four undirected sub-graphs are generated by the filtering algorithm for CS#1.

Figure 3 depicts the filtered undirected matrix and subgraph for the data transmission relationship type for a change in the payload technology (CS#1). A total of twelve sub-graphs were created for this case study using four relationship types for each of three change scenarios.

	Mission_Controller	Joystick	External_Video_Recorder_opt	Video_Digitizer	Converter_Hub	GS_GPS_Receiver	GS_GPS_Antenna	Electronic_Speed_Controller	AV_CDL_Antenna	AV_CDL_Transceiver	AV_PDL_Antenna	AV_PDL_Transmitter	AV_GPS_Receiver	AV_GPS_Antenna	AV_Servo_1	AV_Servo_2	AV_AP_Magnetometer_Opt	AV_AP_PWM_Pins	AV_AP_Serial_Ports	AV_AP_Processor	AV_PS_Multiplexer	GS_PDL_Antenna	GS_PDL_Receiver	GS_CDL_Antenna	GS_CDL_Transceiver	AV_AP_IMU_Thermocouples	AV_AP_IMU_Rate_Gyros	AV_AP_IMU_Accelerometers	AV_AP_IMU_Barometric_Pres	AV_AP_IMU_Static_Pressure	AV_PS_SS_Camera_1	AV_PS_SS_Camera_2
Mission_Controller		1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Joystick	1		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
External_Video_Recorder_opt	0	0		1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Video_Digitizer	0	0	1		1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
Converter_Hub	1	0	0	1		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
GS_GPS_Receiver	1	0	0	0	0		1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
GS_GPS_Antenna	0	0	0	0	0	1		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Electronic_Speed_Controller	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AV_CDL_Antenna	0	0	0	0	0	0	0	0		1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
AV_CDL_Transceiver	0	0	0	0	0	0	0	0	1		0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
AV_PDL_Antenna	0	0	0	0	0	0	0	0	0	0		1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AV_PDL_Transmitter	0	0	0	0	0	0	0	0	0	0	1		0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
AV_GPS_Receiver	0	0	0	0	0	0	0	0	0	0	0	0		1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
AV_GPS_Antenna	0	0	0	0	0	0	0	0	0	0	0	0	1		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AV_Servo_1	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AV_Servo_2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AV_AP_Magnetometer_Opt	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
AV_AP_PWM_Pins	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	1	0		0	1	0	0	0	0	0	0	0	0	0	0	0	0
AV_AP_Serial_Ports	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0		1	0	0	0	0	0	0	0	0	0	0	0	0
AV_AP_Processor	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1		0	0	0	0	0	1	1	1	1	1	0	0
AV_PS_Multiplexer	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0	0	1	1
GS_PDL_Antenna	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0		1	0	0	0	0	0	0	0	0	0
GS_PDL_Receiver	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1		0	0	0	0	0	0	0	0	0
GS_CDL_Antenna	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0		1	0	0	0	0	0	0	0
GS_CDL_Transceiver	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1		0	0	0	0	0	0	0
AV_AP_IMU_Thermocouples	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0		0	0	0	0	0	0
AV_AP_IMU_Rate_Gyros	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0		0	0	0	0	0
AV_AP_IMU_Accelerometers	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0		0	0	0	0
AV_AP_IMU_Barometric_Pres	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0		0	0	0
AV_AP_IMU_Static_Pressure	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0		0	0
AV_PS_SS_Camera_1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0		0
AV_PS_SS_Camera_2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	

Figure 3. Undirected DSM (top) and Network Graph (bottom) for Payload Change Scenario (CS#1) Data Transmission

Additionally, a list of nodes and relationships for each subgraph is generated so that the SME can evaluate each of the included relationships by answering the question "Which direction does the change flow?" for a given change scenario and relationship type. Here, it is important to note that the SME's response is dictated by the level of understanding of the overall system and the individual components. A thorough knowledge of the contingency margins within the system is ideal since the propagation of change is highly dependent on these margins. Then, using the perceived direction of change flow, a directed subgraph is created which will be used to calculate the CPI of each component within the filtered graph. This step allows the SME to explicitly document the perceived direction of change flow and provide reasoning for the elimination of any edge in the filtered graph. For example, in CS #1 the undirected subgraph for power indicates existing data flows from the ground station converter hub (Converter_Hub) to the ground station communications data link (GS_CDL_Transceiver), which then connects all components receiving information from the communications data link; however the SME indicated that the converter hub would effectively shield those components from any change in the payload data transmission since the communications data link operates on a different radio frequency than the payload data link. Therefore, those components were eliminated from the data transmission subgraph for CS #1. Figure 4 displays the directed filtered matrix and subgraph for CS#1 data transmission.

	Mission_Controller	External_Video_Recorder_opt	Video_Digitizer	Converter_Hub	AV_PDL_Antenna	AV_PDL_Transmitter	AV_PS_Multiplexer	GS_PDL_Antenna	GS_PDL_Receiver	AV_PS_SS_Camera_1	AV_PS_SS_Camera_2
Mission_Controller		0	0	1	0	0	0	0	0	0	0
External_Video_Recorder_opt	0		1	0	0	0	0	0	0	0	0
Video_Digitizer	0	0		0	0	0	0	0	1	0	0
Converter_Hub	0	0	1		0	0	0	0	0	0	0
AV_PDL_Antenna	0	0	0	0		1	0	0	0	0	0
AV_PDL_Transmitter	0	0	0	0	0		1	0	0	0	0
AV_PS_Multiplexer	0	0	0	0	0	0		0	0	1	1
GS_PDL_Antenna	0	0	0	0	1	0	0		0	0	0
GS_PDL_Receiver	0	0	0	0	0	0	0	1		0	0
AV_PS_SS_Camera_1	0	0	0	0	0	0	0	0	0		0
AV_PS_SS_Camera_2	0	0	0	0	0	0	0	0	0	0	

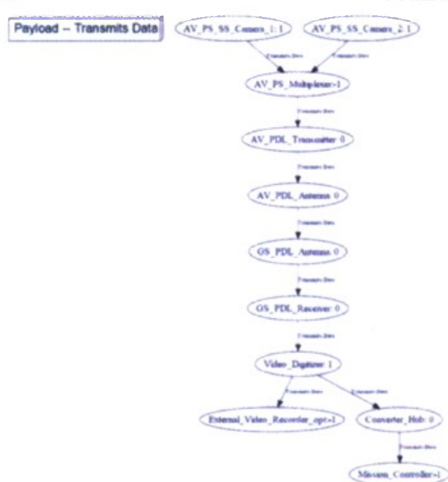

Figure 4. Directed DSM (top) and Subgraph (bottom) for Payload Change Scenario (CS#1) Data Transmission

Multiple change initiators may result in multiple subgraphs for a given relationship type if there is a disconnect in the change flow. For example, CS #2 for the data transmission relationship, the payload data link (AV_Payload_Data_Link and GS_Payload_Data_Link) transmits on a separate frequency than the communications data link (AV_Comm_Data_Link and GS_Comm_Data_Link) so the information flows are independent. The resulting subgraph will contain two change trees that are not connected as shown in Figure 5. Both trees should be considered in the analysis since other relationship types may link components within both change trees.

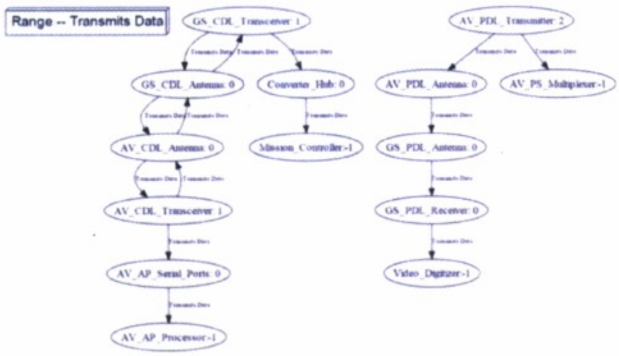

Figure 5. Directed Subgraph for Range Change Scenario (CS#2) Data Tranmission

Identifying Change Multipliers/Carriers. Suh (2005) provides a process to calculate the CPI for each node in the filtered subgraph.

$$\mathrm{CPI}_i = \sum_{j=1}^{n_{out}} \Delta E_{out,j} - \sum_{k=1}^{n_{in}} \Delta E_{in,k}$$

Equation 1. Change Propagation Index

Nodes having a CPI > 0 are change multipliers, indicating that the node propagates more change than it absorbs. A CPI < 0 categorizes the node as a change absorber. While these two types of change behavior are typically the most interesting, change carriers, nodes having a CPI = 0, are also important when attempting to design flexible systems, since they effectively provide a pass through for change. Eckert et al. (2004) acknowledges that system components do not have a predetermined change behavior. Similarly, the CPI as a metric of the change behavior varies across change scenarios and the context in which the system is examined.

Filter by Relationship Type and Change Scenario. The CPI was first calculated for each node within the subgraphs that were filtered by relationship type for each change scenario. To complete the previous example, Figure 6 displays the CPI calculation given filtering for CS#1 data transmission.

	Mission_Controller	External_Video_Recorder_opt	Video_Digitizer	Converter_Hub	AV_PDL_Antenna	AV_PDL_Transmitter	AV_PS_Multiplexer	GS_PDL_Antenna	GS_PDL_Receiver	AV_PS_SS_Camera_1	AV_PS_SS_Camera_2	Ein
Mission_Controller		0	0	1	0	0	0	0	0	0	0	1
External_Video_Recorder_opt	0		1	0	0	0	0	0	0	0	0	1
Video_Digitizer	0	0		0	0	0	0	0	1	0	0	1
Converter_Hub	0	0	1		0	0	0	0	0	0	0	1
AV_PDL_Antenna	0	0	0	0		1	0	0	0	0	0	1
AV_PDL_Transmitter	0	0	0	0	0		1	0	0	0	0	1
AV_PS_Multiplexer	0	0	0	0	0	0		0	0	1	1	2
GS_PDL_Antenna	0	0	0	0	1	0	0		0	0	0	1
GS_PDL_Receiver	0	0	0	0	0	0	0	1		0	0	1
AV_PS_SS_Camera_1	0	0	0	0	0	0	0	0	0		0	0
AV_PS_SS_Camera_2	0	0	0	0	0	0	0	0	0	0		0
Eout	0	0	2	1	1	1	1	1	1	1	1	
CPI	-1	-1	1	0	0	0	-1	0	0	1	1	
Class	A	A	M	C	C	C	A	C	C	M	M	

Figure 6. CPI Calculation for Payload Change Scenario (CS#1) Data Transmission

Filtered by Change Scenario. Eckert (2004) does not distinguish between the relationship type in the classifying of the change behavior (i.e., multiplier or absorber). Thus, an aggregated CPI was calculated by summing component CPI for each relationship type within the change scenario. This result provides an indication of which components are overall multipliers/carriers/absorbers for the change scenario, independent of the relationship type. Figure 7 displays the aggregated CPIs for CS#1 given three relationship types (power, data transmission, houses).

	Mission_Controller	External_Video_Recorder_opt	Video_Digitizer	Converter_Hub	AV_PDL_Antenna	AV_PDL_Transmitter	AV_Power_Switch	AV_Power_Supply	Motor	Electronic_Speed_Controller	AV_AP_Voltage_Current_Regulator	Payload_Pod	AV_PS_Voltage_Regulator	AV_PS_Multiplexer	GS_PDL_Antenna	GS_PDL_Receiver	AV_PS_SS_Camera_1	AV_PS_SS_Camera_2	Ein
Mission_Controller		0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
External_Video_Recorder_opt	0		1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Video_Digitizer	0	1		1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	3
Converter_Hub	1	0	1		0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
AV_PDL_Antenna	0	0	0	0		1	0	0	0	0	0	0	0	0	0	0	0	0	1
AV_PDL_Transmitter	0	0	0	0	1		0	0	0	0	0	0	0	1	0	0	0	0	2
AV_Power_Switch	0	0	0	0	0	0		1	0	0	0	0	0	0	0	0	0	0	1
AV_Power_Supply	0	0	0	0	0	0	0		0	0	0	0	1	0	0	0	0	0	1
Motor	0	0	0	0	0	0	0	0		1	0	0	0	0	0	0	0	0	1
Electronic_Speed_Controller	0	0	0	0	0	0	0	1	0		0	0	0	0	0	0	0	0	1
AV_AP_Voltage_Current_Regula	0	0	0	0	0	0	0	1	0	0		0	0	0	0	0	0	0	1
Payload_Pod	0	0	0	0	0	0	0	0	0	0	0		1	1	0	0	1	1	4
AV_PS_Voltage_Regulator	0	0	0	0	0	0	0	0	0	0	0	0		1	0	0	0	0	1
AV_PS_Multiplexer	0	0	0	0	0	1	0	0	0	0	0	0	0		0	0	2	2	5
GS_PDL_Antenna	0	0	0	0	1	0	0	0	0	0	0	0	0	0		1	0	0	2
GS_PDL_Receiver	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1		0	0	2
AV_PS_SS_Camera_1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0		0	1
AV_PS_SS_Camera_2	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0		1
Eout	1	1	3	2	2	2	0	3	0	1	0	0	2	5	1	2	3	3	
CPI	0	0	0	0	1	0	-1	2	-1	0	-1	-4	1	0	-1	0	2	2	
Class	C	C	C	C	M	C	A	M	A	C	A	A	M	C	A	C	M	M	

Figure 7. CPI Calculation for Payload Change Scenario (CS#1) Given Three Relationship Types

Filtered by Relationship Type. Another context for CPI analysis may include examining the change behaviors of the nodes for each relationship type across all change scenarios. A calculation of CPI given this context would require summing the CPI for each node in all of the data transmission subgraphs (ie data transmission subgraphs for CS#1, CS#2, and CS#3). Figure 8 displays the CPI calculation filtered only by the relationship type data transmission.

	Mission Controller	External_Video_Recorder_opt	Video_Digitizer	Converter_Hub	AV_CDL_Antenna	AV_CDL_Transceiver	AV_PDL_Antenna	AV_PDL_Transmitter	AV_AP_Serial_Ports	AV_AP_Processor	AV_PS_Multiplexer	GS_PDL_Antenna	GS_PDL_Receiver	GS_CDL_Antenna	GS_CDL_Transceiver	AV_PS_SS_Camera_1	AV_PS_SS_Camera_2	Ein
Mission Controller		0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	2
External_Video_Recorder_opt	0		1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Video_Digitizer	0	1		1	0	0	0	0	0	0	0	0	2	0	0	0	0	4
Converter_Hub	1	0	1		0	0	0	0	0	0	0	0	0	0	1	0	0	3
AV_CDL_Antenna	0	0	0	0		1	0	0	0	0	0	0	0	0	1	0	0	2
AV_CDL_Transceiver	0	0	0	0	1		0	0	0	0	0	0	0	0	0	0	0	1
AV_PDL_Antenna	0	0	0	0	0	0		2	0	0	0	0	0	0	0	0	0	2
AV_PDL_Transmitter	0	0	0	0	0	0	2		0	0	1	0	0	0	0	0	0	3
AV_AP_Serial_Ports	0	0	0	0	0	1	0	0		0	0	0	0	0	0	0	0	1
AV_AP_Processor	0	0	0	0	0	0	0	0	1		0	0	0	0	0	0	0	1
AV_PS_Multiplexer	0	0	0	0	0	0	0	2	0	0		0	0	0	0	1	1	4
GS_PDL_Antenna	0	0	0	0	0	0	2	0	0	0	0	1	0	0	0	0	0	3
GS_PDL_Receiver	0	0	1	0	0	0	0	0	0	0	0	2		0	0	0	0	3
GS_CDL_Antenna	0	0	0	0	1	0	0	0	0	0	0	0	0		1	0	0	2
GS_CDL_Transceiver	0	0	0	0	0	0	0	0	0	0	0	0	0	1		0	0	1
AV_PS_SS_Camera_1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0		0	1
AV_PS_SS_Camera_2	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0		1
Eout	1	1	3	3	2	2	4	4	1	0	3	3	2	1	3	1	1	
CPI	-1	0	-1	0	0	1	2	1	0	-1	-1	0	-1	-1	2	0	0	
Class	A	C	A	C	C	M	M	M	C	A	A	C	A	A	M	C	C	

Figure 8. CPI Calculation for Relationship Type Data Transmission Given All Three Change Scenarios

Formulate Real Options. Using the CPI calculations above, the designer can begin to formulate real options for flexible design. Suh (2005) suggests looking first for change multipliers or components that the multipliers are propagating change to, and adjusting their contingency margins to absorb the change, thus making the system more robust to the change scenario. Similarly, designers could seek out change multipliers to locate where it might be good to embed flexibility in order to change the propagation paths. In the event that the change multipliers cannot be modified or if the change propagation paths are not altered by the possible modifications, then an option to modify surrounding change carriers may provide the desired outcome. While this approach to creating options is dependent on the goal/benefit of exploiting the options (i.e., to reduce the cost impact of the change on the system or to allow for the system to adopt new capabilities in the future), utilizing the results of the CPA can assist designers in: 1) identifying potential locations for options and 2) providing an improved estimate of the switch costs associated with the option due to the inclusion of the change paths. Once real options have been formulated, they can then be valued using ROA tools. In some cases, the switch cost of implementing an option may be more than the perceived benefit of the changed design when compared to other designs, and this will be observed in the valuation of such options. Wilds et al. (2007) provides an example of the valuation of real options "in" a system using three different ROA tools.

Application of Filtered Out Degree Method

The above discussion of CPA took a bottom-up approach to the problem of flexible system design. CPA focused on the structure and connectivity of system components and then attempted to determine beneficial modification to the structure to limit the impact (and/or take advantage of the upside) of uncertainty on future system performance and cost. Designers can also considers flexibility of the system architecture as a whole. The filtered-out-degree measure described above takes this approach comparing different solution of the MAV system as whole and measure the apparent changeability of the variants. To examine the usefulness of such a metric, a filtered-OD analysis was conducted on the MAV system. Only the results for the first change scenario (CS#1) will be presented here. Analysis for the other scenarios was conducted in a similar fashion.

To compute the filtered-OD, a system model was developed using the Multi-Attribute Tradespace Exploration technique (Ross, 2008). This resulted in a set of 438 distinct designs of the aircraft (different geometries as well internal components). These designs were assessed with respect to their cost and key performance attributes as identified by the decision maker[1] (DM). The DM identified three attributes, aircraft range, endurance and payload capability, as the most important is choosing design. The tradespace contained designs that included a simple payload (daytime imaging only) as well as those that included an advanced (Day and night capable) payload.

CS#1 is further detailed such that initially only the day-capable payload is available and that later the day-night capable may become available. Initial production will be for 50 units, followed by another 100 units once the day-night payload becomes available. Should the new payload be adopted, the initial production lot will be upgraded to include it. There is uncertainty as to the DM preferences with respect to range and endurance once the new payload becomes available.

Given this refined scenario, filtered-OD provides a mechanism for identifying day-capable design that can be upgraded to large variety of day-night capable design. The filtered-OD is defined as the number of day-night capable design a particular day-only design can be transformed subject to a constraint on cost of transformation. This information is valuable to the designer who may have a good sense of the DM's cost but is uncertain as to the performance requirements given the new payload – the designer knows that they will need to change the design, but lack sufficient information to determine which changes are most valuable to the DM.

---

[1] The former government program manager for the system on which to exercise was based served as the proxy for the decision maker.

Designs with greater filtered-OD at a given cost have greater apparent flexibility[2]. An OD function is also defined by varying the cost threshold for permissible transitions. The resulting graph (Figure 9) comparing OD functions of the day-capable designs aides in visualizing the trade-off between flexibility and cost to achieve that flexibility.

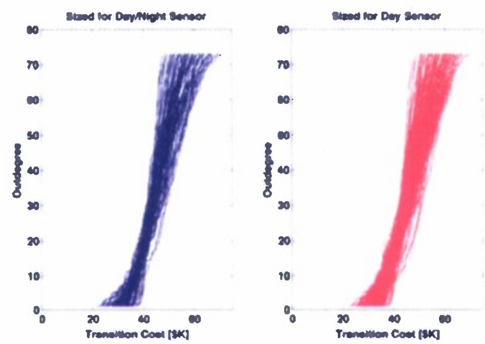

Figure 9. Outdegree as Function of Transition Cost for Payload Change Scenario

Should the designers find the observed OD function are of too high a cost or do not produce sufficient variety of transition possibilities, they can investigate embedding real options into the designs. These options serve as path enablers reducing the cost of transition thereby increasing the f-OD. Of course, having these options may require additional initial expense, but that may be justified by the additional flexibility. As an example consider an alternative design of the day-capable MAV that includes a larger payload bay to ease upgrade to the day-night capable MAV. Figure 10 shows the OD functions for this set of designs compared to those that only considered the day imaging payload. Note the reduction of transformation cost for given f-OD.

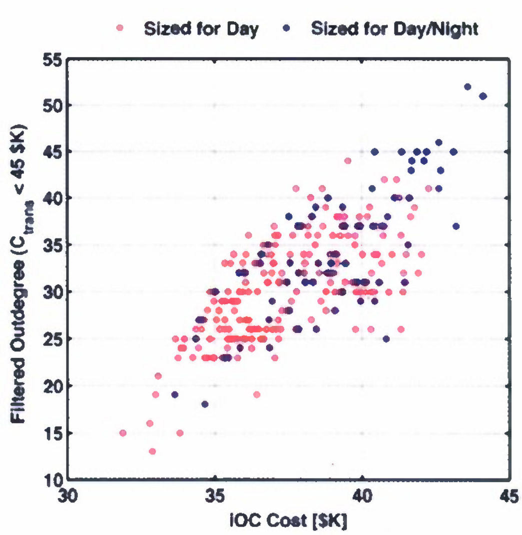

Figure 10. OD Functions for Payload Change Scenario

[2] The term flexibility is used here instead of the more general changeability because the particular type of transition under consideration involves an intervention by an external agent to change the payload.

Valuing Flexibility Using Real Options Analysis

Real options analysis tools can be used to value options at predetermined time to suggest potential decision paths. These decision paths are strongly influenced by uncertainty models and the ability to estimate the associated switch costs. Three types of real options analysis tools include Net Present Value (NPV) calculations, Decision Analysis and binomial lattice methods.

NPV is the total of the present values of all future amounts, typically representing the total yearly profits discounted to present value. Discounting typically involves reducing future profits by a discount rate, or factor, on an annualized basis to account for the fact that a dollar tomorrow is worth less than a dollar today. The calculation is very sensitive to the chosen discount rate and the discount rate is typically treated as a constant over the period of analysis. The discount rate is often very difficult to determine from a political perspective, due to adjustment for perceived risks, though not difficult to calculate from a technical perspective. Discount rates are often set by high-level management or decision makers. (de Neufville 1990)

Decision Analysis (DA) accounts for the value of flexibility by structuring possible contingent decisions in a decision tree. Designers then choose the solution that offers the best expected value, a weighted average of the outcomes by their probability of occurrence (de Neufville, 1990). Wang (2005) points out that the expected value is based on an NPV calculation, and thus suffers from the same challenge of determining and using a fixed discount rate.

Binomial lattice method is based on a collapsed representation (i.e., later states are a multiple of earlier states) of the evolution of contingent decisions. This method assumes path independence and requires knowledge of the volatility and predicted growth rates of the modeled uncertainties. However, this information is not readily available for new and emerging technologies due to the absence of historical data. Furthermore, the lattice method assumes a constant expected growth rate, which is not typical of newly emerging technologies.

Wilds et al. (2007) uses NPV, DA, and binomial lattice method to provide an example of how to apply real options "in" a design. The paper includes a comparative assessment of the three ROA tools noted in this paper using a single MAV case study with comparable real options to those discussed in this paper. Wilds et al. (2007) conveys the importance of considering the assumptions of each method, with respect to known and unknown information, in order to carefully select the best ROA tool for valuing options.

Conclusions

CPA, as extended from the cited work, allows the designer to investigate how possible changes will impact the structure and behavior of a system design. The method outlined can aid designers in identifying system components that have some likelihood of changing. Designers can focus efforts on reducing change costs and impacts by embedding options into the design. Using the multipliers and carriers as guides, designers can develop potential options for inclusion, and then use ROA to determine which options generate sufficient benefits to justify costs. CPA does have some limitations, however. CPA results are highly sensitive to the particular set of change scenarios considered. If changes cannot be well represented or characterized, the analyst may miss identifying the change initiator or incorrectly propagate the change. CPA also relies upon representation of a system as a static graph of interactions between components (the DSM). Complex systems that change structure or behavior in response to changing contexts, i.e. self-modifying or intelligent systems, may not be easily represented using such a construct. Furthermore, traditionally DSMs only represent binary relationships between components. In complex systems, there may be multiple types of relationships between components. For example, two components may be physically connected, as well as exchange electrical energy. In different change scenarios, one type of relationship may result in change propagation, while the other does not. System representation that explicitly includes multiple relationship types, and filtering by those types when analyzing changes, helps to mitigate this issue. However, dealing with change scenarios that involve multiple relationship types is an ongoing research challenge.

CPA and f-OD can be used in concert to increase flexibility at a system level. Designers can use change scenarios to motivate system transition options or paths, and use f-OD to find system designs that are more flexible. The OD function can provide decision makers with a visual representation of the tradeoff between cost incurred in exercising transitions and the variety of transitions available from which to choose. Since the cost of transition is directly related to the changes in the system that occur during a transition, CPA can be used to determine where in the system those costs are being incurred, and to identify portions of the system that could benefit from redesign (e.g. through the addition of options) to reduce transition costs and/or to increase the variety of transitions available at a given cost. Taken together, CPA and f-OD can be used to help guide designers to generate and place real options to enable valuably flexible systems.

References

Bartolomei, Jason, Capt USAF, "Multi-Design Optimization Analysis for Endurance vs. Longest Linear Dimension." Technical Report, Massachusetts Institute of Technology, 2005.

Bartolomei, Jason, Capt UASF, *EPLANE_MAV.xls* USAF Academy, 2005.

Clarkson, J.P., Simons, C., and Eckert, C., "Predicting Change Propagation in Complex Design." *Technical Proceedings of the ASME 2001 Design Enginineering Technical Conferences* (Pittsburgh, PA, Sep 9-21, 2001), DETC2001/DTM-21698.

de Neufville, R., *Applied Systems Analysis*. McGraw-Hill, New York, NY, 1990.

Eckert, C., Clarkson, J.P., and Zanker, W., "Change and Customisation in Complex Engineering Domains." *Research in Engineering Design,* 15(1):1-21, 2004.

Ross, A.M., Rhodes, D.H., and Hastings, D.E., "Defining Changeability: Reconciling Flexibility, Adaptability, Scalability, Modifiability, and Robustness for Maintaining Lifecycle Value." *Journal of Systems Engineering*, 11(3), 2008 forthcoming. (Preprint available at http://seari.mit.edu)

Shah, N.B., Viscito, L., et al., "Quantifying Flexibility for Architecting Changeable

Systems." *6th Conference on Systems Engineering Research* (Los Angeles, CA, April 4-5, 2008), University of Southern California, Paper #154.

Suh, E.S., "Flexible Product Platforms." Doctoral Dissertation, Massachusetts Institute of Technology, Engineering Systems Division, Cambridge, MA, 2005.

Wang, T., "Real Options 'in' Projects and Systems Design—Identification of Options and Solution for Path Dependency." Doctoral Dissertation, Massachusetts Institute of Technology, Engineering Systems Division, Cambridge, MA, 2005.

Wilds, J. W., Bartolomei et al., "Real Options In a Mini-UAV System." *5th Conference on Systems Engineering Research, Track 1.1: Doctoral Research in System Engineering (*Hoboken, NJ, March 14-16, 2007), Stevens Institute of Technology, Paper #112.

Designing a Perception-Based Decision Support System to Improve System Acquisition Decisions

M.L. Cummings

Anna E. Bethke

Abstract:

For system acquisition decisions, decision makers utilize multiple criteria and many levels of reasoning to understand complex multivariate data. These decision makers could be greatly aided by direct-perception decision support systems, specifically configural as compared to traditional separable displays. Given both designs, an experiment was conducted that compared performance across different levels of reasoning. The configural display promoted better performance and more efficient eye fixation patterns at the highest level of reasoning better than the separable display, and was also the subjective tool of choice. Design implications for future system acquisition decision support systems are discussed.

1 Introduction

System acquisitions are a crucial component of military operations. The source selection milestone is a critical step, yet is often the most difficult portion of an acquisition since it requires decision makers to objectively understand large-scale system trade-offs through analysis of complex, multivariate sets of quantitative data. Analyzing this large, multiple criteria set of data is difficult, underscoring the importance of providing decision makers with an intuitive depiction of this information. The lack of decision support for a decision maker, who is faced with a complex decision based on multiple criteria, presents a significant research gap. To this end, this research focuses on creating a decision support system which displays system acquisition information in a more intuitive, principled format.

2 Design Requirements

In the development of a proof-of-concept decision support system for system acquisition decisions, we focused on the following three fundamental metrics, which are core to all acquisition decisions [1]: 1) The degree to which pre-specified functional requirements are met, 2) The degree to which desired non-functional requirements (commonly known as "-ilities") are met, and 3) Costs, which can be decomposed across system lifecycle or functional requirements. While there are other factors to be considered in such decision which will be discussed later such as risk and schedule, these three criteria encapsulate the primary drivers of source selection, i.e., does the candidate system meet an organization's needs and what will it cost?

Given these attributes of system acquisition source selection decisions, the ultimate goal of a decision maker is to answer feasibility questions (i.e., does a system meets a set of selection criteria?), or optimality questions (i.e., in the case of deciding between competitive systems, which system best meets the desired criteria?). Highlighting the subjective nature of such decisions, the selection criteria could emphasize cost over function, or could emphasize that all proposed functional requirements be met, regardless of cost. Thus, these decisions are inherently cost-benefit trade-based, so any effective decision support tool with allow a decision maker needs to make comparisons within and across system metrics.

We propose that to be able to conduct relevant trade space analyses, decision makers incorporate three general levels of reasoning for these complex acquisition decisions. They are:

1) Data Processing: Low-level reasoning that compares values within a single constraint. This requires a simple search of a data set for an answer, for example, determining which competitive system has the overall greatest cost.
2) Information Aggregation: Mid-level reasoning that integrates data across a single constraint. For example, determining which system meets a minimum level of specified functional requirements requires decision makers to integrate multiple values of similar data types.
3) Knowledge Synthesis: High-level of reasoning that requires the integration of information across multiple constraints [2]. For example, determining which system has the lowest cost, meets all "-ilities," and meets all functional requirements requires decision makers to integrate information across multiple data sets and perform some optimization to determine the best case.

The current state-of-the-art in terms of system acquisition decision support generally includes spreadsheets and more classical graphical representation such as bar and line charts. These displays are known as seperable displays because these charts typically represent each state variable in a singular fashion [3]. However, as illustrated by the three different levels of reasoning, decision makers typically must integrate various sources of information. Thus, we propose that a better systems acquisition tool is one that promotes data integration. This hypothesis will be explored in the next section.

3 Fan Visualization

Because system acquisition decisions require both the integration and comparison of information, we propose that a configural display could best support these types of decisions. Configural displays integrate multivariate information as compared to separable displays, which assign unique representations to each state variable [3]. An example of a configural display is a radar chart while an example of a separable display is a dial gauge. A configural display maps individual variables in such a way to create emergent features, so that users directly perceive relationships within the data [4] [5]. As a result, often these displays take on some kind of geometrical representations. Configural displays have been shown to improve a user's performance while completing data integration problems [6, 7].

Leveraging these ideas of direct-perception, we developed a decision support tool, Fan Visualization, (FanVis), for system acquisition decision support tool which consists of a series of configural displays. The following trade space variables were included to support cost-benefit analyses: 1) The degree to which functional requirements are met (f.r. met), 2) The degree to which "-ilities" are met ("-ilities" met), 3) Total cost, 4) Cost per sub-functional requirement (Cost per sub-f.r.), and 5) Life cycle cost.

In total, there are 5 two-dimensional views in FanVis: 1) the System View, 2) the Multi-System View, 3) the Functional Requirement View, 4) the Comparison View, and 5) the "-ility" View. All the views were built upon the System View to provide the decision maker with different perspectives of the acquisition trade space, and are described in detail in the next sections. Details of the software architecture can be found in [8].

3.1 The System View

The main structure of FanVis is similar to a radar chart where variables (in this case, the functional requirements of the proposed system) are represented by axes which originate from a central point. Each system in the design space is represented by a polygon in the System View, such as in Figure 1. The vertices of a candidate system's polygonal representation intersect the pre-specified functional requirement axes at particular points along those axes to demonstrate how well the system meets each particular functional requirement. The axes scales are a five-point Likert scale [9] with the following delineations: 1) Does not Meet Requirements (closest to the central point), 2) Partially Meets Requirements, 3) Meets Requirements (middle point, shown in red), 4) Exceeds Requirements, and 5) Greatly Exceeds Requirements (furthest from the central point). Faint lines connect the axes along this five point scale to provide a visual anchor. This type of scaling was selected since humans generally cannot assign meaningful absolute weights to categorical variables beyond five gradations [10].

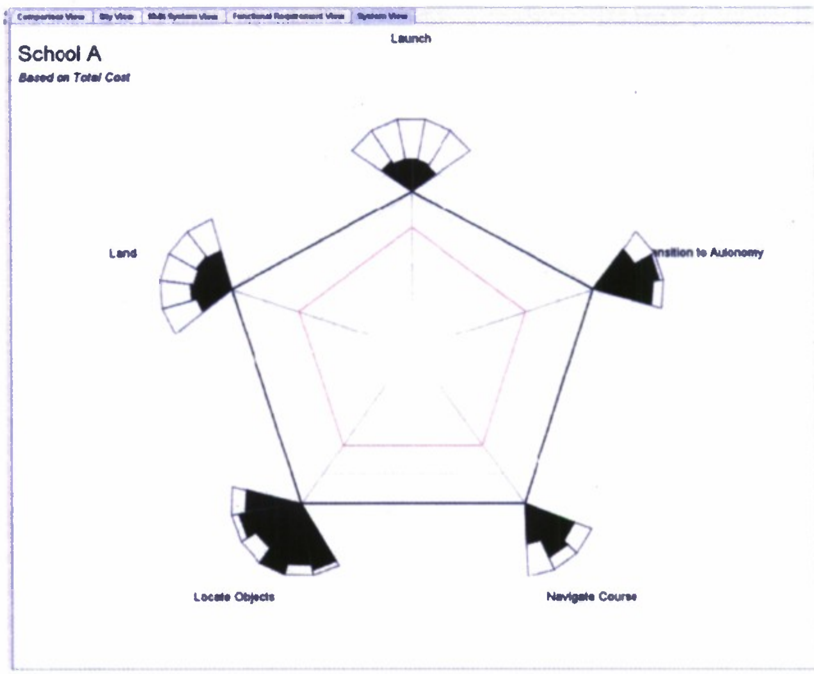

Figure 1: System View of FanVis

At the terminus of each outer axis is a fan comprised of blades, which represents the first layer of sub-functional requirements. Each fan not only encodes the degree to which the associated functional requirement is met, but it also represents the overall cost of the functional requirement and the cost of the sub-functional requirements within the fan via shading. In this manner a single data element encodes three trade space variables. If there is a functional requirement that is driving the cost of the system, the fan representing that requirement will be mostly black, while all other fans will be mostly white. By having one fan different than all other fans, it will be more salient, thus giving the association that the difference should be noticed and potentially remedied. If the functional requirements are balanced in cost, all the fans will be mostly black.

3.2 Multi-System View

The Multi-System View displays two or more system views side by side, as shown in Figure 2, promoting direct comparisons among competing systems. Because all necessary information is positioned within the user's visual scan, the decision maker benefits from uninterrupted visual reasoning [11], which allows decision makers to focus on the differences between the systems easily, without having to switch views and integrate information from disparate sources. For instance, Figure 2 quickly reveals that the system at the right is less capable than those on the left as illustrated by its smaller size, but the cost distribution is very similar. This type of direct comparison could greatly help the decision makers conduct their cost-benefit analyses.

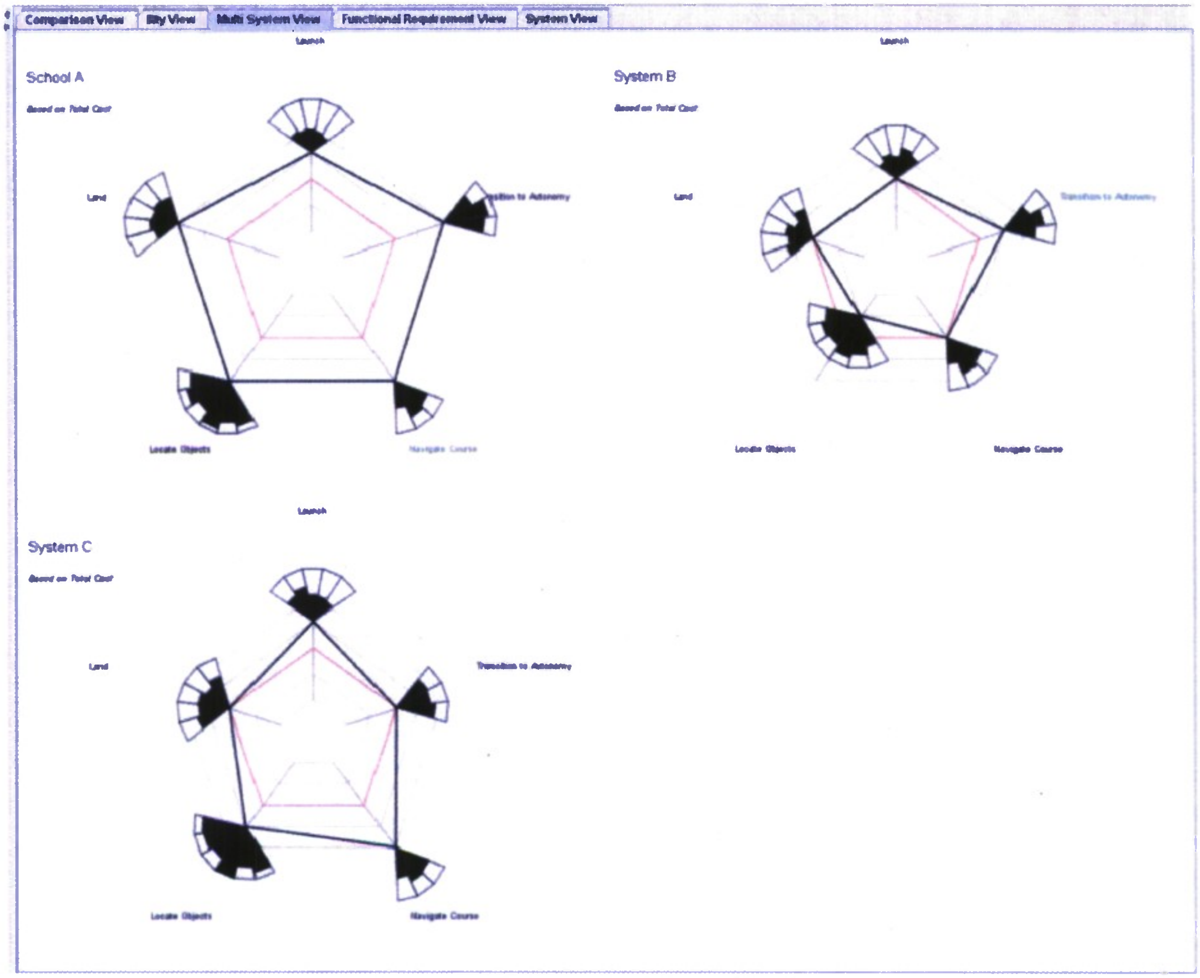

Figure 2: Multi-System View of FanVis

3.3 Functional Requirement View

The Functional Requirement View (Figure 3) displays multiple systems for a single functional requirement. This view allows decision makers to probe deeper into the potential tradeoffs within the trade space. It permits users to view multiple systems, as in the Multi-System View, but allows a greater degree of detail. In this view, the degree to which the functional requirements are met has been modified from the polygonal structure to flat lines with the delineations: 1) Does not Meet Requirements (bottom line), 2) Partially Meets Requirements, 3) Meets Requirements (middle line, shown in red), 4) Exceeds Requirements, and 5) Greatly Exceeds Requirements (top line). For example, in Figure 3 the first system greatly exceeds the requirement, the second only partially meets the requirement, and the third exceeds the requirement.

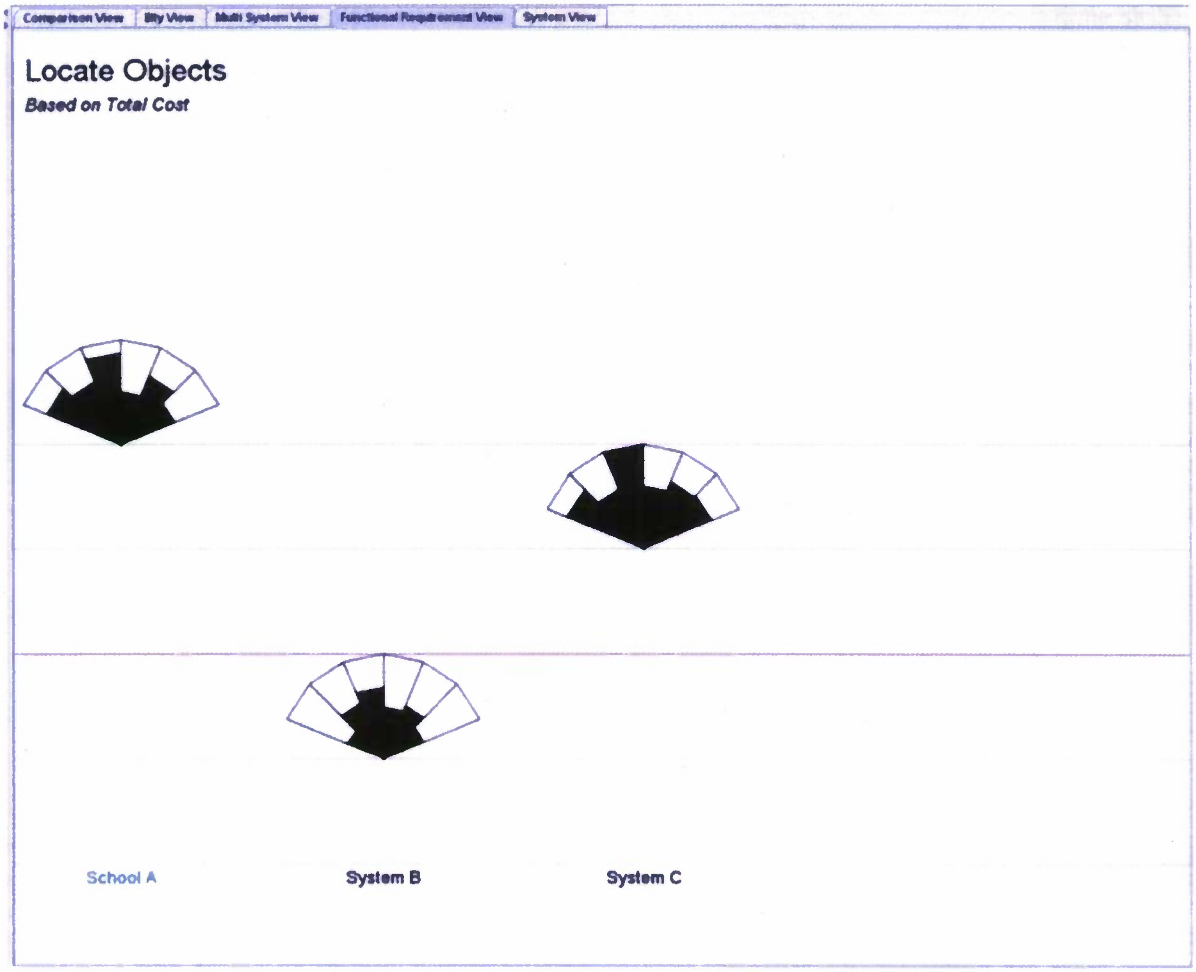

Figure 3: Functional Requirement View of FanVis

3.4 Comparison View

The Comparison View, as shown in Figure 4 provides a higher level of data abstraction by displaying two or more systems without the lower level sub-functional requirement cost information. This gives the user the ability to obtain a global view of the trade space. Each polygon represents a single system; however, the fans have been removed, deleting information regarding the sub-functional requirement costs. Instead, total cost is displayed as a function of the color of the system's polygonal representation. The color of each system polygon is relative to the other systems in the trade space. A color legend in the lower left of Figure 4 displays both relative systems' costs as well as digital values. The color gradient is an interval sequence which ranges from blue for the most expensive system, to yellow for the least expensive system.

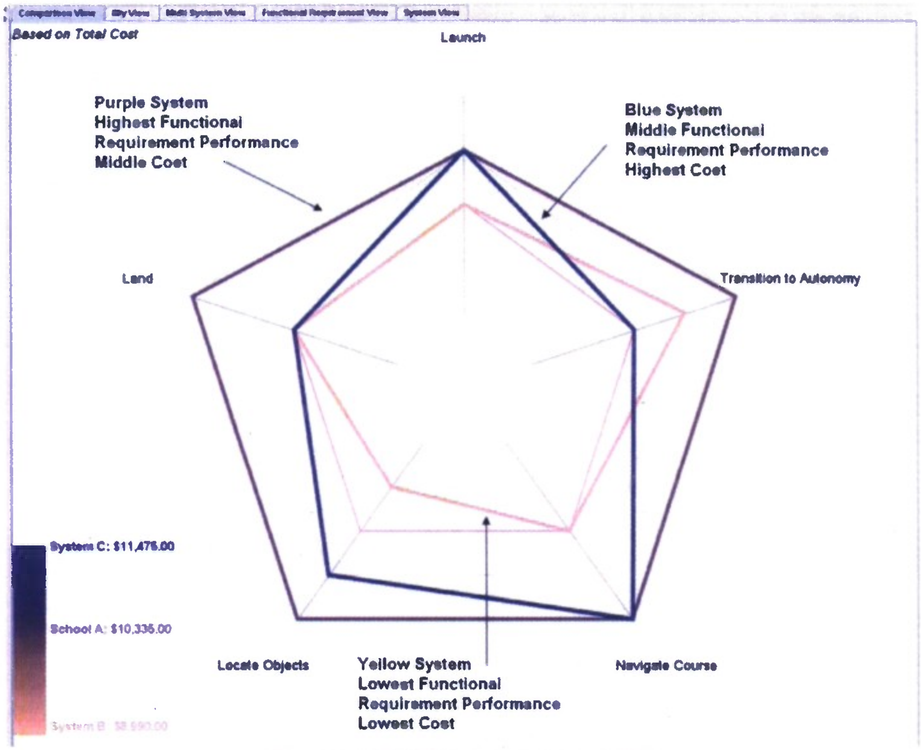

Figure 4: FanVis Comparison View

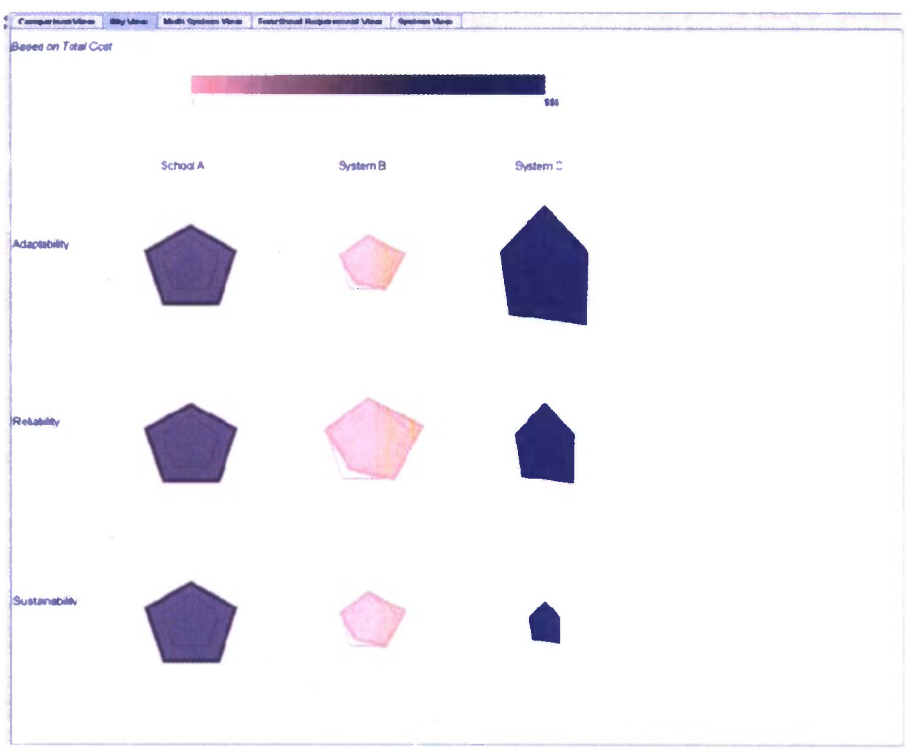

Figure 5: FanVis "-ility" View

3.5 The "-ility" View

Decision makers can obtain additional information in the "-ility" View, as shown in Figure 5. After entering in those "-ilities" from a menu that he or she deems important (e.g., maintainability, survivability, etc.), a decision maker can analyze how well "-ilities" are met in addition to analyzing the degree to which the requirements and cost are met. How well each "-ility" is met is shown by scaling the size of a system's polygonal representation. This scale is a three-point Likert scale ranging from 1) Does not Meet Requirements (shrinking the polygon from its original size), 2) Meets Requirements (original size) and 3) Exceeds Requirements (expanding the polygon from its original size). A three point Likert scale is used instead of the five point Likert scale for this view due to the subjective nature of "-ilities." For the most part, "-ilities" are quality attributes of a system with less clearly defined threshold performance criteria, so measuring them on a finer scale could ultimately lead to data misconceptions by the decision makers [1]. In this manner, a smaller polygon represents that a system that falls short of some desired level of the selected "-ility", so the better system would have large polygons for each "-ility."

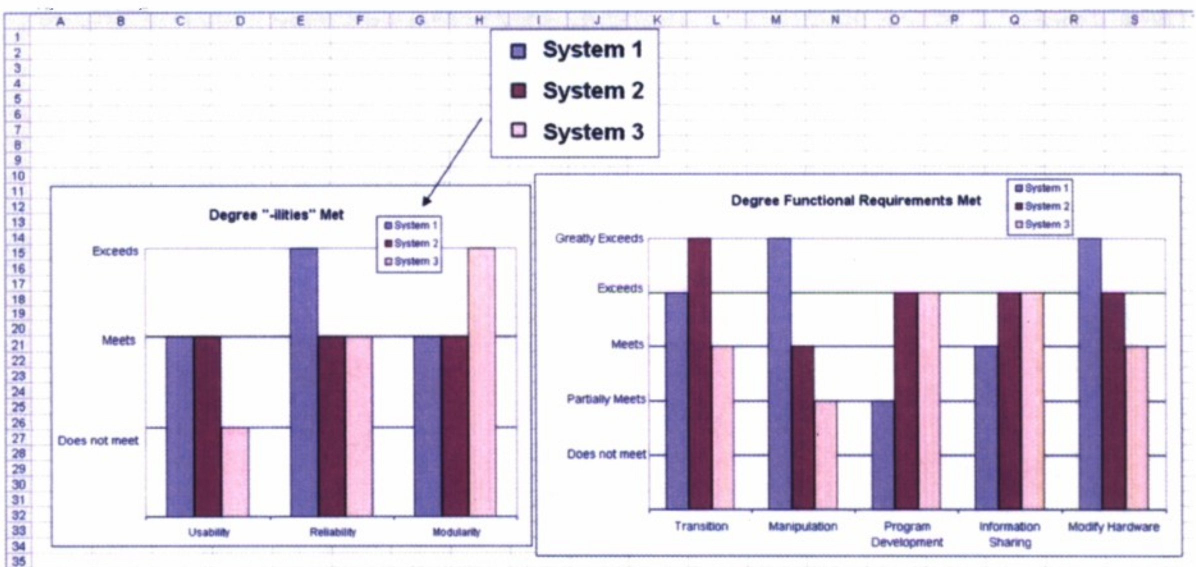

Figure 6: Requirements and "-ilities" Tab of Excel®

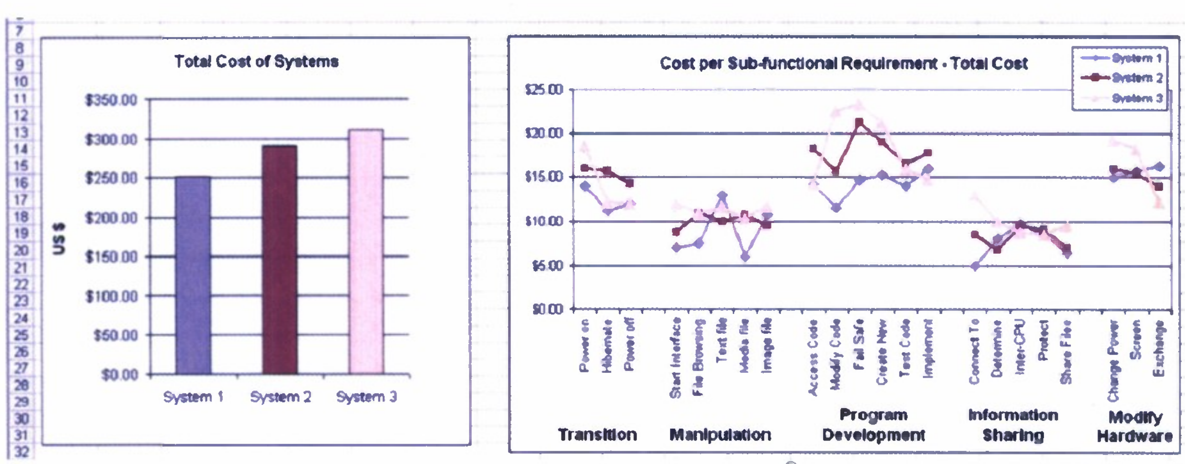

Figure 7: Total Cost Tab of Excel®

3.6 Separable Decision Support Tool Design

In order to determine if FanVis promotes superior decision performance, a more traditional separable decision support tool was also created to represent the current state of the art. The separable tool was created in Excel® since it is a very common application for conducting system acquisition decisions. The tool utilizes four tabs which mimic the views provided in FanVis. Three of the tabs, Requirements and "-ilities", Total Cost, and Cost Categories, are graphical displays of the data, while the last tab, Data, includes the raw numbers of the trade space.

The Requirements and "-ilities" tab displays the degree to which each requirement is met by systems in the trade space in two bar charts. Each requirement or "-ility" is represented by a different bar, while each system has a different color code (Figure 6). This color scheme was the Excel® default for three variables in a line chart. This color scheme was retained as it gave sufficient separation among the three colors.

Unlike FanVis where data representations were often encoded from multiple trade space variables, each data element within the Excel® tool only represents a single trade space variable. For example, in Figure 6, the data elements in the left bar chart each represent how a system meets an "-ility." The data elements in the right bar chart each represent how a system meets a functional requirement. There are no data elements which represent how the functional requirements as a whole are met. Instead, decision makers must integrate this information.

The Total Cost Tab displays each system's total cost in a bar chart as well as the cost per sub-functional requirement in a line chart, as shown in Figure 7. A line connects the costs of the sub-functional requirements within a given functional requirement. These lines help the decision maker delineate the various functional requirements from each other for a given system. The Cost Category Tab is much like the Total Cost Tab. For each life cycle cost phase under consideration, the total system cost appears for all systems in a bar chart. In addition, the cost per sub-functional requirement for all systems is shown in line charts by cost phase.

Lastly, the Data Tab is simply the raw trade space numbers, organized by system. All trade space data can be found under this tab including the degree to which a system meets the requirements and "-ilities", as well as the cost per sub-functional requirement for each of the life cycle phases under consideration. The raw data was provided in the separable tool as users can access and manipulate this data in FanVis through a tree selection menu [8].

4 Experimental Evaluations

It was hypothesized that the configural decision support tool, FanVis, would be able to support high-level system acquisition decisions to a greater degree than the traditional separable decision support tool. This hypothesis was tested in terms of not just participant performance on system acquisition tasks, but also subjective appeal of the decision support tools.

4.1 Participants

To test these hypotheses, 30 participants between the age of 18 and 75 were recruited for this study. The average participant age was 52.4 years with a standard deviation of 11.4 years. Personnel with experience in either high-level system acquisitions or high-level organizational were specifically recruited for this experiment as both roles utilize the high-level data analysis skills required for a system acquisition and source selection. All experiment participants had moderate to high levels of experience completing high-level decisions for a team, project or organization, and 23 of the 30 had detailed system acquisition experience. The average number of years of system acquisition experience of those participants was 14.0 years with a standard deviation of 13.1 years. All had experience using data manipulation tools such as Excel®, and none indicated that they were color blind.

Table 1: Experimental System Aequisition Seleetion Criteria

- At least meet all "-ilities"
- At least meet all functional requirements
- Minimize cost
- Maximize degree "-ilities" and functional requirements met
- Balance
    - "-ilities" across system
    - Functional requirements across system
    - Cost across sub-functional requirements

4.2 Experimental Procedure

The experiment eonsisted of seven parts: pre-experiment interaetions, a baseline data handling profieieney test, two training sessions, two test sessions, and post-experiment debriefing. On average, the full experiment lasted an hour and a half. The experimental tasks were performed on the lower three displays of a six-display, Windows-based workstation mounted inside a 2006 Dodge Sprinter Van known as the Mobile Advaneed Command and Control Station (MACCS).

After signing a eonsent form and filling out a brief demographie survey, partieipants took a baseline data handling profieieney test to assess their Exeel® familiarity and data processing skills. This test was eonstrueted from the Edueational Testing Serviee® praetiee questions for the quantitative seetion of the Graduate Reeord Examinations [12, 13], and eonsisted of quantitative multiple ehoiee questions that were answered by interpreting Exeel® eharts and graphs. The partieipants' baseline data handling profieieney was based on the time to answer all questions.

The partieipants were given a tutorial of the deeision support tool and a representative ease study, in whieh they were to seleet the best eandidate system. The system seleetion eriteria given to partieipants in deereasing order of importanee are listed in Table 1. The tutorial gave partieipants an overview of eaeh deeision support tool, how data was eneoded within the tool, and speeifie features that would likely be neeessary to utilize while eompleting the test session. During the tutorial, the partieipants were able to see and interaet with both tools using a praetiee data set and were eneouraged to ask questions. Partieipants were eneouraged to praetiee using both tools until they felt eomfortable with their use. Following the tutorial, the eye traeker, an ISCAN® Polhemus VisionTraek® system [14] was ealibrated.

For the FanVis test session, a ease study regarding the seleetion of a student Unmanned Aerial Vehiele (UAV) system by a funding ageney was presented, while in the Exeel® tool, a ease study regarding laptop seleetion by a board member of a low-ineome sehool distriet was used. These two ease studies were built from the same trade spaee data in order to ensure that the two test sessions were similar in diffieulty. However, the laptop data set was sealed down by 3.5% so the numbers would not appear exact. The UAV study was an actual contest so there was a known ground truth in terms of the best system aetually selected.

In each test session, the partieipant answered questions regarding the system acquisition trade spaee deseribed in the ease study. The partieipants were asked 19 identieal (in format and diffieulty) trade spaee questions for both tools. These questions began with six Knowledge Synthesis questions, followed by six Information Aggregation questions, six Data Proeessing questions and eoneluded with a repeat of an initial Knowledge Synthesis question. The first and last question was "Whieh system best meets the baseline system seleetion eriteria?" The last question was repeated to determine if the exploration of the data spaee, held eonstant for everyone, ehanged the partieipants' final deeisions. All questions had a definitive eorreet answer. For eaeh question, four ehoiees were presented to the partieipant.

Eaeh partieipant eompleted two test sessions; one with Exeel® and the other with FanVis. The order of these test sessions was eounterbalaneed and randomized. A brief retrospeetive protoeol was conducted following completion of both test sessions in order to determine why a participant manipulated the tools in a specific manner, as well as subjective responses in terms of acceptance and preference.

4.3 Experimental Design

The experiment was a 2x3 repeated measures design with two independent variables: Decision Support Tool (FanVis vs. Excel®) and Reasoning Difficulty Level (Knowledge Synthesis, Information Aggregation, Data Processing). All participants received all six treatment combinations. The order that the participants received the two levels of Decision Support Tool was counterbalanced and randomly assigned. The Reasoning Difficulty Level was presented in the same order for all participants to more accurately reflect actual practices of current decision makers who typically start with broad, more ambiguous questions, and then drilled down through hierarchical levels of information to obtain answers. Moreover, this was held constant to ensure an approximate similar exploration strategy of the trade space. Dependent variables included decision time per Reasoning Difficulty Level (Knowledge Synthesis, Information Aggregation, Data Processing), Reasoning Difficulty Level accuracy score (%), fixation data which will be described in detail below, and lastly, subjective preference questions.

5 Results

For statistical analysis, a 2x3 repeated measures mixed linear model was applied to analyze the decision time dependent variable, Time. For all other dependent variables, non-parametric Mann Whitney and Wilcoxon-Signed Rank tests were used since parametric assumptions were not met. An alpha level of 0.05 for all statistical tests was used.

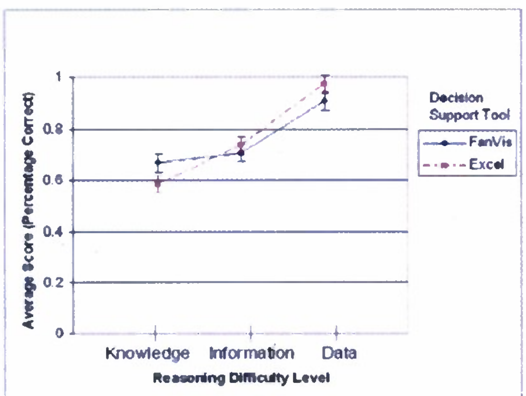

Figure 8: Average percentage of correct answers

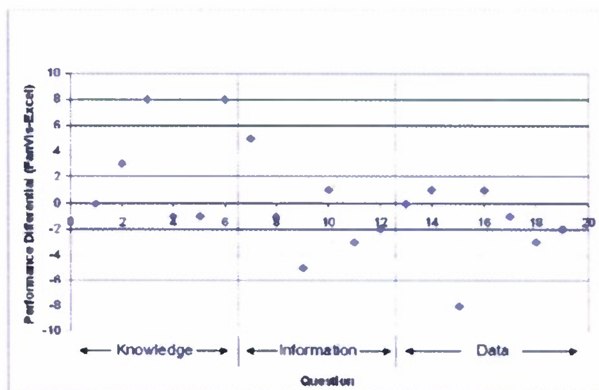

Figure 9: Total performance differential by question

5.1 Reasoning Difficulty Level Accuracy Score

As expected, a participant's accuracy score varied depending upon the reasoning difficulty level. As shown in Figure 8, participants obtained a higher score using FanVis when answering Knowledge Synthesis questions (Mann-Whitney Dependent Test, $z=1.99$, $p=0.046$), while participants obtained a higher score using Excel® while answering Data Processing questions ($z=2.21$, $p=0.027$). There was no statistical difference between the two decision support tools for Information Aggregation score ($z=0.77$, $p=0.437$).

Each participant's performance was then analyzed on a per-question basis to determine when a participant answered a specific question correctly with one tool and incorrectly with the other. Because an identical set of questions was asked for both tools, this metric indicates if a participant was only able to extract the required information out of one of the two tools. For instance, if a participant answered a question correctly while using Excel® but the same question incorrectly using FanVis, then there is a performance differential in favor of Excel®. The total performance differential by question is shown in

Figure 9. Of interest to note are those questions which have a large differential. Questions 3, 6 and 7 have particularly high performance differentials in favor of FanVis. These questions all asked the participant to integrate the cost of the sub-functional requirements. On the other hand, questions 9 and 15 have particularly high performance differentials in favor of Excel®. These two questions asked the participants to extract the "-ility" data. It appears from these results that FanVis promoted better sub-functional requirement analysis, while Excel® was better for "ility" analysis.

Furthermore, the participants' performance differential on the question "Which system best meets the baseline system selection criteria?" (The first and last question asked) was analyzed to determine if exploring the data space changed the participants' decisions given the different systems. In FanVis, 2 participants answered the question incorrectly both times, 1 incorrectly changed his initial answer, 6 correctly changed their initial answers, and 21 answered correctly both times. In all, a marginally statistically significant portion of the participants changed their answers ($z=-1.890$, $p=0.059$), suggesting interaction with FanVis allowed the participants to obtain a clearer picture of the system acquisition trade space over the time frame of the experiment.

5.2 Decision Time

A logarithmic transformation (natural log) of the time dependent variable was utilized to satisfy normality and homogeneity assumptions. In addition, two outliers (greater than 3.29 standard deviations from the mean) were deleted from the data set [15]. This 2x3 mixed linear model included proficiency time as a covariate. Proficiency time, as described previously, was the time it took the participants to answer a set of baseline data processing questions based on Excel® charts and graphs.

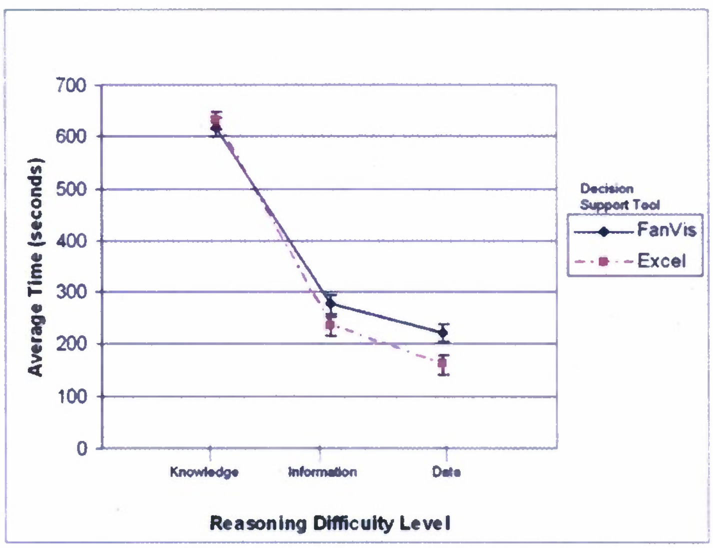

Figure 10: Average time to answer questions

Given this model, a significant difference was found for Decision Support Tool ($F(1,29)= 12.17$, $p=0.0016$) and Reasoning Difficulty Level ($F(1,29)=216.71$, $p<0.001$). There was no interaction effect. Figure 10 yields insight into the actual average time participants spent to answer the system acquisition decision support questions. For the relevant pair wise comparisons across the Reasoning Difficulty Levels, only the Data Processing level pair was significantly different ($p<0.001$). Participants spent an average of 60.8 seconds more time answering these questions in FanVis compared to Excel®. It is likely that this additional time in FanVis occurred due to the greater number of data elements within the tool. FanVis contained a total of 278 data elements (defined as elements which directly encode data as opposed to elements which manipulate said data), whereas the Excel® tool contained 171 data elements. Accessing these additional hundred elements could have caused this time disparity.

5.3 Fixation Patterns

Analyzing the fixation patterns of the participants for the two tools can yield insight into the participant's cognitive strategies [16, 17]. This section describes fixation patterns that were common within the two tools, and postulates the similarities and differences caused by the two different tools, as well as their implications. Unfortunately, there was a considerable amount of noise present in the eye tracker data, which is common for these devices [18] . For this reason, only five participants (3, 6, 16, 23, and 24) had eye tracks which were continuously accurate for the duration of both the FanVis and Excel® test sessions.

While there was no clear difference between FanVis and Excel® in terms of efficient fixation patterns, there were differences in the time that participants fixated on relevant elements in the two systems. A higher percentage of time spent fixating on relevant elements indicates participants spent a greater amount of time fixating on data elements useful to answering the questions [18]. Overall, participants spent an average of 66.18% of the time fixating on relevant elements in FanVis compared to 66.58% in Excel®, which is not statistically different. However, analyzing the percent time fixating on relevant elements by Reasoning Difficulty Level yields interesting trends. Participants fixated on relevant elements in FanVis 75.81% of the time for Knowledge Synthesis questions as compared to 65.80% of the time while using Excel®, which is a significant difference (Wilcoxon-Signed Rank Test, $z=-2.023$, $p=0.043$). This trend is reversed for Data Processing questions where participants had a slightly higher percentage time fixating on relevant elements in Excel® (65.59%) compared to FanVis (64.75%), though this comparison was not statistically significant ($z=-0.730$, $p=0.465$). There was no statistical difference between the two tools for the information aggregation level fixation times. This trend mirrors the performance results in that FanVis was more useful for knowledge synthesis, while Excel® was more helpful for answering data processing questions.

5.4 Subjective Responses

Overall, participants felt both tools allowed them to complete the necessary tasks, and many commented either tool was a vast improvement on how they currently complete acquisitions, which often requires sifting through reams of paper. In terms of tool preference, a statistically significant portion of the participants felt FanVis was a more useful tool than Excel® (Mann-Whitney Dependent test, $z=2.01$, $p=0.04$). They also felt that FanVis was able to give them a better understanding of the trade space than Excel® ($z=3.10$, $p=0.002$), and would choose to use FanVis over Excel® given the opportunity during their next system acquisition decision ($z=2.01$, $p=0.04$). A marginally statistically significant portion of the participants felt FanVis was a more pleasant tool to use than Excel® ($z=1.64$, $p=0.1$).

Overall, twenty-six percent of the participants felt that given time, they would be able to find data more effectively in FanVis. Thirteen percent of the participants commented that FanVis was more intuitive, thirteen percent commented they enjoyed being able to dig deeper into the data in FanVis, and thirteen percent commented FanVis was more visually appealing. On the other hand twenty percent of the participants commented it was easier for them to view the bar charts in Excel® and determine how the requirements were being met versus using FanVis.

6 Conclusions

Because system acquisition decisions require knowledge-based reasoning through the integration and comparison of multiple criteria, multivariate information, we proposed that a configural versus a separable display would best support typical decision-makers in these settings. Given the design of a configural display that leveraged perceptual-based reasoning (FanVis), the results from an experiment suggest that a configural tool better supports system acquisition decisions as compared to a more traditional spreadsheet Excel®-based separable display. Given the fact that system acquisition decisions typically involve three levels of reasoning (data processing, information aggregation, and knowledge synthesis), Table 2 summarizes the experimental results. FanVis was clearly the superior display for support of knowledge-based questions, while the Excel® display better supported users in determining low-level data processing answers. The displays were equivalent across the information aggregation level of reasoning.

Table 2: Summary of Experimental Findings

	Knowledge Synthesis	Information Aggregation	Data Processing
Score	FanVis ($p=0.046$)	No Difference ($p=0.437$)	Excel® ($p=0.027$)
Time	No Difference ($p=0.471$)	No Difference ($p=0.032$)	Excel® ($p<0.001$)
% Time Fixating on Relevant Elements	FanVis ($p=0.043$)	No Difference ($p=0.50$)	No Difference ($p=0.46$)
Subjective Opinion	FanVis ($p=0.040$)		

These results are important for two reasons. First, while all system acquisition decisions require low level reasoning, ultimate source selection decisions are inherently knowledge-based, so given a choice between a decision support system that better supports data processing reasoning or knowledge synthesis, the system that promotes knowledge should be the more desirable one. However, this choice is not mutually exclusive in that these results also clearly demonstrate that a hybrid configural-separable display would most likely be the best choice for actual implementation.

The second most important result in Table 2 is the fact that participants felt FanVis was a more useful tool. Moreover, as discussed earlier, users generally felt FanVis gave them a better understanding of the trade space, and would choose to use it in their next acquisition decision. The overall preference for FanVis is a particularly striking result since participants all had extensive experience using Excel®, while none had ever seen or used FanVis prior to the experiment. Often familiarity with a decision support tool biases users to this tool, and can leave them wary of a revolutionary tool (e.g., [19]). Overall, participant preference for FanVis after only about 90 minutes of interaction, coupled with the fact that users were all very experienced and familiar with Excel®-based spreadsheets, demonstrates that such use of a configural display for multiple criteria decision making has significant potential.

Acknowledgements

The Air Force Office of Scientific Research (AFOSR) sponsored this research and the Office of Naval Research sponsored the eye tracker.

References

1. Blanehard, B.S. and W.J. Fabryeky, *Systems Engineering and Analysis*. 1998, Upper Saddle River, NJ: Prentiee Hall.
2. Rasmussen, J., *Skills, Rules, and Knowledge; Signals, Signs, and Symbols, and Other Distractions in Human Performance Models*. lEEE Transaetions on Systems, Man, and Cyberneties, 1983. SMC-13(3): p. 257-266.
3. Bennett, K.B. and M.L. Toms, *Emergent Features and Graphical Elements: Designing More Effective Configural Displays*. Human Faetors, 1993. 35: p. 71-97.
4. Gibson, J.J., *The ecological approach to visual perception*. 1979, Boston, MA: Houghton Mifflin.
5. Hansen, J.P., *An Experimental Investigation of Configural, Digital, and Temporal Information on Process Displays*. Human Factors, 1995. 37(3): p. 539-552.
6. Carswell, M.C. and C.D. Wickens, *Mixing and Matching Lower-Level Codes for Object Displays: Evidence for Two Sources of Proximity Compatibility*. Human Faetors, 1996. 38(1): p. 1-22.
7. Elvers, G.C. and N.J. Dolan, *A comparison of the augmented bar display and the object display*. Ergonomies, 1995. 38(4): p. 777 - 792.
8. Massie, A.E., *Designing a Graphical Decision Support Tool to Imporve System Acquisition Decisions*, in *MIT Aeronautics and Astronautics*. 2009: Cambridge, MA.
9. Likert, R., *A technique for the measurement of attitudes*. Arehives of Psychology, 1932. 140.
10. MeCormick, E.J., *Human Factors in Engineering and Design*. 4th ed. 1976, New York: MeGraw-Hill Book Co.
11. Tufte, E.R., *Envisioning Information*. 1990, Chesire, CT: Graphies Press.
12. ETS. *Graduate Record Examinations General Test Practice Book*. 2007 [eited 2007 February]; Available from: http://www.ets.org/Media/Tests/GRE/pdf/grc_0809_practiee_book.pdf.
13. ETS. *Math Review for Practicing to Take the GRE(R) General Test*. 2007 [cited 2007 February]; Available from: http://www.ets.org/Media/Tests/GRE/pdf/GREmathPractice.pdf.
14. lne., I. *ISCAN Eye & Traget Tracking Instumentation*. [cited April 21, 2009; Availablc from: http://www.iscaninc.com.
15. Tabaehniek, B.G. and L.S. Fidell, *Using Multivariate Statistics*. 4th ed. 2001, New York: HarperCollins. 880.
16. Hornof, A.J. and T. Halverson, *Cognitive Strategies and Eye Movements for Searching Hierarchical Computer Displays*, in *Proceedings of the SIGCHI conference on Human factors in computing systems*. 2003: Ft. Lauderdale, Florida, USA.
17. Goldberg, J.H. and X.P. Kotval, *Computer Interface Evaluation Using Eye Movements: Methods and Constructs*. International Journal of lndustrial Ergonomics, 1999. 24: p. 631-645.
18. Poolc, A. and L.J. Ball, *Eye Tracking in Human-Computer Interaction and Usability Research: Current Status and Future Prospects*, in *Encyclopedia of Human Computer Interaction*, C. Gaoui, Editor. 2005, lGI Global: Hcrshcy, PA.
19. Bruni, S. and M.L. Cummings. *Human Interaction with Mission Planning Search Algorithms*. in *ASNE Human Systems Integration Conference*. 2005. Arlington, VA.